ΑD	1

Award Number: DAMD17-99-1-9135

TITLE: Characterization of Two C. Elegans Homologuses of Oncogenic Inhibitor of Apoptosis Proteins (IAPs) and Identification of Interacting Genes

PRINCIPAL INVESTIGATOR: Andrew G. Fraser, Ph.D. Dr. Julie Ahringer

CONTRACTING ORGANIZATION: University of Cambridge
Cambridge CB2 1TS England

REPORT DATE: September 2000

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

# REPORT DOCUMENTATION PAGE

Form Approved OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Artington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY	USE	ONLY	(Leave
hlank)			

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED

September 2000

Annual Summary (1 Aug 99 - 1 Aug 00)

#### 4. TITLE AND SUBTITLE

Characterization of Two C. elegans Homologuses of Oncogenic Inhibitor of Apoptosis Proteins (IAPs) and Indentification of Interacting Genes

5. FUNDING NUMBERS DAMD17-99-1-9135

#### 6. AUTHOR(S)

Andrew Fraser, Ph.D.

## 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Cambridge

8. PERFORMING ORGANIZATION REPORT NUMBER

#### E-MAIL:

Agflo@hermes.cam

## 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

#### 11. SUPPLEMENTARY NOTES

This report contains colored photos

#### 12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

## 13. ABSTRACT (Maximum 200 Words)

Inhibitor of Apoptosis Proteins (IAPs) suppress apoptosis both in the fruit fly Drosophila melanogaster and in vertebrates. All IAPs contain at least a single copy of a highly conserved domain, the Baculovirus IAP Repeat (BIR) domain; this is essential for their anti-apoptotic activity. Understanding how a BIR domain regulates apoptosis is thus an important step in furthering our understanding of the molecular mechanisms of apoptosis. I have previously identified two BIR-containing Proteins (BIRPs) in the nematode worm C. elegans. One of these, BIR-1, appears to play no role in the regulation of programmed cell death in C. elegans; however, BIR-1 is required for the completion of cytokinesis. Cytokinesis is a complex process that is likely to be highly regulated and to involve a large number of proteins. My approach to understanding how BIR-1 functions in cytokinesis has been to try to identify other genes that are also required for cytokinesis in C. elegans and subsequently to attempt to understand how BIR-1 relates to this machinery. Rather than examine individual candidate genes to determine whether they have a role in cytokinesis, I wished to carry out a genome-wide screen to identify all genes that are required for cytokinesis; to do this I have made use of RNA-mediated inhibition (RNAi). RNAi is a technique whereby the activity of a particular gene is transiently inhibited following the introduction of dsRNA of sequence specific to the targeted gene. The specificity and potency of RNAi make it an ideal technique to investigate gene function beginning only with genomic sequence. I have screened ~90% of genes on Chromosome I of C. elegans for RNAi phenotypes and have identified 5 genes which have a similar RNAi phenotype to BIR-1. I aim to extend this to appraoch over the entire genome in the coming year and thence to spend time characterising genes found to be required for cytokinesis.

14. SUBJECT TERMS

Breast Cancer

15. NUMBER OF PAGES

52

17. SECURITY CLASSIFICATION OF REPORT

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

16. PRICE CODE

Unclassified

Unclassified

Unclassified

Unlimited

## **Background**

Many tumour cells have lost the ability to undergo apoptosis in response to DNA damage caused either by irradiation or chemical mutagens. This appears to be one of the key reasons that so many tumours are resistant to treatment by either radio- or chemotherapy. Understanding how apoptosis is regulated in normal cells may shed light on the mutations that render tumour cells resistant to apoptotic stimuli and may thus suggest novel therapeutic strategies. The key focus of this project is to understand how a particular conserved family of regulators of apoptosis, the Inhibitor of Apoptosis Proteins (IAPs), suppresses cell death.

IAPs suppress apoptosis both in the fruit fly *Drosophila melanogaster* and in vertebrates. All IAPs contain at least a single copy of a highly conserved domain, the Baculovirus IAP Repeat (BIR) domain; this is essential for their anti-apoptotic activity. Understanding how a BIR domain regulates apoptosis is thus an important step in furthering our understanding of the molecular mechanisms of apoptosis.

The nematode worm *C. elegans* has been a key tool for research into apoptosis since the inception of the field. Genetic analysis of programmed cell death in *C. elegans* led to the identification of the basal machinery of cell death which is conserved between worms and humans. Analysis of the function of IAPs in *C. elegans* might shed light on the function of IAPs in human cells. I have previously identified two BIR-containing Proteins (BIRPs) in the nematode worm *C. elegans*. One of these, BIR-1, appears to play no role in the regulation of programmed cell death in *C. elegans*; however, BIR-1 is required for the completion of cytokinesis. Furthermore, I demonstrated that a human homologue of BIR-1, the BIRP survivin, can partially substitute for BIR-1 in the nematode; this shows that BIRPs have a conserved role in the regulation of cytokinesis. These results have subsequently been extended by other groups who have shown that both fission and budding yeasts contain BIRPs and that in

each case inhibition of these genes leads to profound cell cycle defects and polyploidy.

The BIR domain thus seems to have a role in cytokinesis in eukaryotic cells ranging from single-celled yeasts to human cells. My focus has been to use *C. elegans* to understand the function BIRPs in cytokinesis: how they are regulated, what precise functional role they have, and which proteins do they interact with.

# **Experimental Approach**

Cytokinesis is a complex process that is likely to be highly regulated and to involve a large number of proteins. My approach to understanding how BIR-1 functions in cytokinesis has been to try to identify other genes that are also required for cytokinesis in *C. elegans*. In this way I hope to be able to build up a more comprehensive view of the cytokinesis machinery and subsequently to attempt to understand how BIR-1 relates to this machinery. This approach has recently been used very productively by the lab of Bob Horvitz, who has shown that inhibition of Aurora-like kinase activity leads to a very similar defect in cytokinesis to that seen following inhibition of BIR-1 and, following this observation, that BIR-1 appears required for the localization of Aurora-like kinase to the cytokinesis furrow and mid-body.

Rather than examine individual candidate genes to determine whether they have a role in cytokinesis, I wished to carry out a genome-wide screen to identify all genes that are required for cytokinesis; to do this I have made use of RNA-mediated inhibition (RNAi). RNAi is a technique whereby the activity of a particular gene is transiently inhibited following the introduction of dsRNA of sequence specific to the targeted gene. The specificity and potency of RNAi make it an ideal technique to investigate gene function beginning only with genomic sequence. Ingestion of dsRNA-expressing bacteria results in RNAi of the targeted gene and we have established that this technique is at least as effective as the injection of dsRNA for RNAi. It is thus possible to make a library of bacteria, each expressing dsRNA corresponding to an individual gene, to target each and every predicted gene in the *C. elegans* genome.

I have made such a library for all genes on Chromosome I (~13% of all predicted genes) and have screened these genes for RNAi phenotypes. I have identified 339 genes with RNAi phenotypes of which 221 are embryonic lethal (as would be expected for a gene with a defect in cytokinesis). This analysis of chromosome I is the first systematic reverse genetic analysis of a multicellular organism and has resulted in a Nature Article.

To further characterise the nature of the defect arising from RNAi of each of the 221 genes that are required for embryonic viability, our lab is currently in the process of making time-lapse movies of embryos for all 221 embryonic lethal genes to determine the nature of the defect that gives rise to embryonic lethality. Thus far we have identified several genes that are required for cytokinesis. These include profilin, an actin-binding protein; a gene encoding a protein with high homology to a centromeric protein INCENP; and a homologue of the *S. cerevisiae* SCD6 gene which may be involved in vesicle fusion and trafficking, a process thought to be involved in *C. elegans* cytokinesis.

#### **Future Work**

I am in the process of extending the RNAi analysis of the *C. elegans* genome to encompass the remaining five chromosomes. I anticipate that construction of the dsRNA-expressing bacterial library should be complete by early 2001, and that analysis of the RNAi phenotypes of all embryonic lethal genes by time-lapse videomicroscopy should be complete by summer 2001. By the time this screen is completed, I expect to have identified 30-50 genes that are required for cytokinesis. This should have greatly expanded our knowledge of the molecular components of the cytokinesis machinery of *C. elegans*. I will then attempt to elucidate how BIR-1 interacts with the identified gene products and thus to understand the involvement of BIR-1 in cytokinesis.

### **Achievements**

Cloning of 2496 genes to generate a dsRNA-expressing bacterial library.

- Screening the 2496 gene bacterial library to identify RNAi phenotypes of the cloned genes.
- Identification of 339 genes with RNAi phenotypes, of which 221 are required for embryonic viability. Detailed analysis of these data, culminating in publication as a Nature Article (attached).
- Timelapse videomicroscopic analysis of the RNAi phenotypes of embryos for 150 of the 221 genes required for embryonic viability.
- Identification of 5 new genes required for cytokinesis in *C. elegans*.

# Systematic functional genomic analysis of *C. elegans* Chromosome I by RNA interference

Andrew G. Fraser\*<sup>1</sup>, Ravi S. Kamath\*<sup>1</sup>, Peder Zipperlen\*<sup>1</sup>, Maruxa Martinez-Campos\*, Marc Sohrmann† and Julie Ahringer\*.

\*Wellcome/CRC Institute, University of Cambridge, Tennis Court Road, CB2 1QR Cambridge, UK. Phone +44-1223-334144; Fax +44-1223-334089. †The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

1 These authors contributed equally to this work.

Complete genomic sequence is known for two multicellular eukaryotes, the nematode Caenorhabditis elegans and the fruit fly Drosophila melanogaster, and will be soon for humans. However, biological function has been assigned to only a small proportion of the predicted genes in any animal. We used RNAmediated interference (RNAi) to target nearly 90% of predicted genes on C. elegans Chromosome I by feeding worms with bacteria that express double stranded RNA. We have assigned function to 13.9% of the genes analysed, increasing the number of sequenced genes with known phenotypes on Chromosome I from 70 to 378. While most genes with sterile or embryonic lethal RNAi phenotypes are involved in basal cell metabolism, many genes giving post-embryonic phenotypes have conserved but unknown function. In addition, conserved genes are significantly more likely to have an RNAi phenotype than genes with no conservation. We have constructed a reusable library of bacterial clones that permits unlimited future RNAi screens, which should help develop a more complete view of the relationships between the genome, gene function, and the environment.

The complete genomic sequence of an organism is an invaluable tool in understanding the molecular mechanisms underlying its development and function. The nematode worm *C. elegans* is one of two multicellular eukaryotes for which essentially complete genomic sequence is known<sup>1,2</sup>. 36% of predicted *C. elegans* genes have a significant human match<sup>1,3</sup> including many genes implicated in human diseases<sup>3,4</sup>, and functional analysis of the *C. elegans* genome has shed light on many conserved biological processes and molecular pathways. A comprehensive functional analysis of all genes in *C. elegans* would greatly expand our knowledge of conserved gene function. We therefore decided to investigate systematically loss-of-function phenotypes of predicted genes of *C. elegans*, starting with Chromosome I.

RNA-mediated interference (RNAi) is a technique whereby the activity of a gene is transiently inhibited following the introduction of double-stranded RNA (dsRNA) of sequence specific to the targeted gene<sup>5</sup>. The specificity and potency of RNAi make it ideal for investigating gene function beginning only with genomic sequence<sup>6</sup>. Ingestion of dsRNA-expressing bacteria results in RNAi of the targeted gene<sup>7</sup>, and we previously established that this technique is at least as effective as the injection of dsRNA for RNAi<sup>8</sup>: embryonic lethal phenotypes are detected with similar efficiency by feeding and injection, but feeding detects over 50% more postembryonic phenotypes than injection. It is thus possible to make a library of dsRNA-expressing bacteria which could be used for high-throughput genome-wide RNAi screens at very low cost. It is important to note that since RNAi does not efficiently inhibit all genes, an RNAi-based screen will miss some relevant genes. Despite this caveat, RNAi is a useful screening tool to complement classical forward genetics.

## Analysis of the function of genes on Chromosome I by RNAi

We constructed a library of bacteria expressing dsRNA corresponding to genes on Chromosome I. Chromosome I is the second smallest chromosome, has few

duplicated gene clusters and has no striking unusual features<sup>1</sup>. Each individual bacterial clone is able to synthesise dsRNA designed to target a single gene; since gene predictions are still changing, a few primer pairs no longer correspond to single genes (see Methods). In total, the resulting library contains 2445 independent clones, corresponding to 2416 predicted genes, a total of 87.3% of the 2769 currently predicted genes of Chromosome I.

We screened the library to identify genes whose inhibition gives a clearly detectable phenotype in wild-type worms as described in Methods. We were able to assign a phenotype to 13.9% of the analysed genes, raising the number of sequenced genes on chromosome I with known phenotypes from 70 to 378 (Table 1). Many genes have more than one associated phenotype, reflecting that genes frequently have multiple functions in the organism. Furthermore, since we examined worms that were only exposed to dsRNA as larvae or adults as well as their progeny, we could assign postembryonic phenotypes to genes that result in sterility or produce 100% embryonic lethal progeny. A summary of these results and a partial listing of the phenotypes obtained are given in Tables 1 and 3. Full results are in Supplementary Table 1 and arepublicly accessible in WormBase (www.wormbase.org).

Our screen was sufficiently effective to identify 90% of known embryonic lethal genes. In addition, we were able to assign phenotypes to 45% of genes with a known post-embryonic phenotype that should have been detectable in our screen (Table 2 and Supplementary Table 2). However, we failed to find phenotypes for a number of previously characterised genes. In some cases (e. g. *fog-3*), this was not due to an inherent difficulty in inhibiting the genes using RNAi (since we obtained the correct phenotype in a separate experiment), but simply because we overlooked them in the screen. However, only one of eight genes involved in neuronal function gave a detectable RNAi phenotype; this accords well with our finding that neurons appear to be more resistant to RNAi than other cell types<sup>8</sup>. Similarly, we did not detect

phenotypes for several genes involved in sperm development (fer-1, spe-9, and spe-11).

The largest phenotypic class, comprising over 60% of the genes, are those whose inhibition by RNAi gives rise to embryonic lethality, the Emb genes; these include a large number of components of the basal cellular machinery. More interestingly, we find a homologue of the SMN human disease gene<sup>9</sup>, a variety of genes encoding RNA-binding proteins (several such proteins play a role in early polarity; reviewed in <sup>10</sup>), a number of genes involved in chromosome condensation and separation, components of signal transduction pathways and many conserved genes that have no known biochemical function.

The largest class of post-embryonic phenotype is the Uncoordinated (Unc) class. Unc phenotypes arise from defects in the development or function of the neuromuscular system (reviewed in <sup>11</sup>). We find Unc genes encoding proteins involved in vesicle sorting and fusion as well as transcription factors (including a homologue of the zinc finger transcription factor MYT-1 which is only expressed in developing neurons in mammals<sup>12-14</sup>) and components of the cytoskeleton (e. g. a kakapo<sup>15-18</sup> and a talin<sup>19</sup> homologue).

A number of genes showed a high incidence of males (Him) phenotype. *C. elegans* is usually grown as a self-fertilising hermaphrodite with males arising at a low frequency in wild-type cultures due to non-disjunction of the X-chromosome (hermaphrodites have two X chromosomes, males only one). An increased number of males is indicative of either the incorrect segregation and maintenance of chromosomes in the germ line (reviewed in <sup>20</sup>) or defects in sexual specification. The Him genes that we identified include kinesins, a katanin homologue<sup>21,22</sup> and a nuclear hormone receptor.

### Conservation of genes with RNAi phenotypes across eukaryotes

We examined the level of cross-species conservation of the genes for which we detected an RNAi phenotype (Fig 1). To find *C. elegans* genes that are conserved in other species, we identified *C. elegans* genes that have hits with BlastP<sup>23</sup> e-values below 1.00E-06 in *Saccharomyces cerevisiae*, *Drosophila melanogaster* or humans; we define these as a "match". Hits with BlastP e-values below 1.00E-10 and in which the conservation extends over at least 80% of the *C. elegans* protein length, we defined as "homologues"; this category includes orthologues. This provides a conservative estimate of the number of genes with regions of conservation (matches) or homologues, respectively.

We found that genes with RNAi phenotypes were much more likely to have a match (p<0.001) compared to all genes (Fig 1). Most striking is the similarity that we see between *C. elegans* and *Drosophila*: while 42% of *C. elegans* genes have a match and 19% have a homologue in *Drosophila*, we find that over 72% of genes with an RNAi phenotype have a *Drosophila* match and 43% have a homologue (Fig 1). This analysis shows that genes with a required function in *C. elegans* have been highly conserved across eukaryotic evolution. We also find that highly conserved genes are more likely to have an RNAi phenotype than genes that show no conservation: 26% of *C. elegans* genes that have a homologue in one of the organisms examined give an RNAi phenotype compared to only 5% of genes with no conservation (p<0.001).

# Physical distribution on chromosome I of genes with RNAi phenotypes

Genes for which we identified an RNAi phenotype are evenly distributed across the chromosome with the exception of two regions (corresponding to segments 2 and 8-9 in Fig 2a) for which there appears to be a drop in number (p<0.1). These two regions correspond to the two regions of chromosome I that contain locally duplicated gene clusters<sup>1</sup>. We suggest that the reduction in the number of phenotypes observed by RNAi in these regions may be due to gene duplication and thus redundancy of function. It is worth noting that some of the predicted genes in the

duplicated regions may not be expressed: while genes with RNAi phenotypes are equally likely to have an EST in all regions of the genome (see below), there is a significant drop (p<0.05) in the proportion of total genes with ESTs in the second locally duplicated gene cluster region (Fig 2b; 39% of genes in the second cluster have an EST compared with 53% over the entire chromosome). We suggest that a portion of the predicted genes in such regions of duplication may in fact be pseudogenes.

Genes that give RNAi phenotypes are much more likely to have an EST than genes on chromosome I in general (82% versus 53% respectively, p<0.001; Fig 2b). The relatively high percentage of genes with RNAi phenotypes that have ESTs may reflect that these genes are expressed at higher levels. It may also be that many genes that currently lack ESTs are only expressed conditionally; we are unlikely to have found phenotypes for such genes.

In *C. elegans*, there is evidence of differences between the chromosome arms and the central regions (the clusters), suggesting that there might be differences in gene type or function across the chromosome<sup>24</sup>. In general, the distribution of genes in any given phenotypic class was similar to that for all genes with an RNAi phenotype (e. g. Emb genes; compare Fig 2c with 2a). However, genes with viable post-embryonic phenotypes (Pep genes) — those that gave a post-embryonic phenotype without any embryonic or post-embryonic lethality, sterility, or developmental delay — show a trend toward enrichment at the arms of chromosome I (p<0.1). It has been suggested that the chromosome arms may be more prone to mutation and recombination than the central core portion<sup>24</sup> and, if so, that novel gene functions are more likely to evolve in such regions. Our finding that genes which uniquely affect post-embryonic development cluster at the arms supports this model.

Relationships between the predicted biochemical function of a gene product and its RNAi phenotype

To explore the relationship between the biochemical function of a gene product and its mutant phenotype, we categorised the sterile (Ste), embryonic lethal (Emb), uncoordinated (Unc) and viable post-embryonic phenotype (Pep) genes into the functional classes shown in Fig 3a.

Unsurprisingly, genes involved in basal metabolic processes account for ~50% of Ste and Emb genes (Fig. 3a); this confirms that these basic biochemical processes are indeed essential for viability. In contrast, under 20% of Unc and Pep genes encode components of the basal metabolic machinery, whereas more than twice as many encode proteins with more specialized functions (Figs. 3a, b). There is thus a clear difference between the types of gene required for germline function or embryonic viability (which mainly require basal machinery) and those involved in later developmental processes which appear to require proteins either of more specialized functions or of as yet unknown function (Fig. 3b).

A second clear trend is that the number of genes of unknown function increases greatly in the Unc and Pep genes, making this the largest overall class for those phenotypes (Fig. 3). This shift underlies the fact that while we know a great deal about basic metabolic processes of eukaryotic cells (and thus can readily ascribe function to a large proportion of Ste and Emb genes), much is still to be learnt about the complex processes and the genes that regulate the development and function of a multicellular eukaryote. A significant number (~25%) of genes of unknown function have close homologues in *Drosophila* or humans; further study of these may shed light on conserved processes specific to animals.

### Comparison of genes essential for viability of S. cerevisiae and C. elegans

S. cerevisiae was the first eukaryote to be completely sequenced<sup>25</sup> and reverse genetics has been used extensively to investigate S. cerevisiae gene function. In a set of 3680 genes knocked out by targeted disruption, 890 affect viability<sup>26</sup>; we compared

these genes to those that gave different RNAi phenotypes in *C. elegans*. Yeast and worm genes important for viability have a similar distribution within the different functional classes, but are different from the Unc or Pep distributions (Fig 3c; also compare to 3a and 3b). This suggests that similar types of gene are required for viability of yeast and animal cells. A striking difference (p<0.001) is that only ~1% of the genes required for viability in yeast are transcription factors, whereas for *C. elegans* it is ~4% (a similar percentage of the genomes of yeast<sup>27</sup> and *C. elegans*<sup>2</sup> encode transcription factors, 3.3% and 2.5% respectively). This suggests that a large fraction of the *C. elegans* transcription factors required for viability may be involved in specific developmental processes.

## An estimate of the size of the functionally non-redundant genome

What do our data tell us about the size of the functionally non-redundant genome? We screened 12.7% of the *C. elegans* genome and found that 339 genes gave a clearly discernible phenotype. Taking into account the sensitivity of our screen and scaling up to the entire genome, we estimate that ~5400 genes will be individually required for wild-type *C. elegans* development under standard laboratory conditions (~2300 genes for embryonic viability and ~3100 post-embryonically; see Methods for calculation). This is comparable to previous estimates based on forward genetics<sup>28</sup>. We expect that phenotypes for other genes will be identified under novel conditions (e. g. environmental stress), in other genetic backgrounds, or using more refined and restricted screening conditions.

### **Discussion**

We have taken a systematic approach to identify functions for the predicted genes of *C. elegans* Chromosome I. This is the first large-scale reverse genetic analysis of a multicellular organism and has increased by five-fold the number of sequenced genes with known phenotypes on this chromosome.

While we have identified RNAi phenotypes for many genes, some will have eluded our screen for one of at least two reasons. Firstly, RNAi may have been ineffective against the targeted gene. RNAi does not accurately phenocopy the null phenotype of all genes (e. g. genes involved in neuronal function), and may result in either partial or no loss of function. It should also be noted that if multiple genes have regions of identical or near-identical nucleotide sequence, RNAi could target them simultaneously, so that the observed phenotype may be the result of the inhibition of more than one gene. Secondly, we will not have detected either subtle or conditional phenotypes. However, we anticipate that future RNAi-based screens using specific assays should be able to detect phenotypes for many more genes, thus increasing our understanding of C. elegans and hence of metazoan biology in general. Since our library consists of bacterial clones that can be replicated, and the feeding protocol is relatively simple compared with injection, the library can be used repeatedly at low cost and high efficiency for such screens. In addition, we expect that a feeding library and database of associated phenotypes will prove valuable for the positional cloning of genes; currently there are over 300 genes on chromosome I identified by mutation but not yet cloned.

Although the time needed for an RNAi screen using our bacterial library is similar to that for a classical genetic screen, the two approaches have different advantages and will yield different results. Both approaches can be used to screen the entire genome for genes involved in a particular process, and both may identify complete or partial loss-of-function phenotypes. Classical forward genetics generates stable mutant lines that can be maintained indefinitely; furthermore, while some genes are resistant to RNAi, all genes are sensitive to mutagens (albeit to a greater or lesser degree) and could thus be cloned using a classical screen. Also, some mutants isolated by forward genetics are due to gain-of-function mutations, which cannot be generated by RNAi. However, the positional cloning of a gene is often slow and laborious. RNAi, while having the disadvantages mentioned above, has the key advantage of all

reverse genetics: the sequence of the gene is already known, and thus any mutant phenotype observed is automatically connected to a known sequence.

In the future, we aim to extend our library construction and functional analysis to the entire *C. elegans* genome and anticipate that the possibility of genome-wide RNAi screening, in conjunction with other functional genomics approaches such as expression analyses using microarrays<sup>29</sup> and two-hybrid experiments<sup>30</sup> will accelerate *C. elegans* research.

#### Methods

Generation and cloning of PCR products. PCR products were synthesised using BioTaq polymerase (Bioline) in a reaction containing 25ng of *C. elegans* genomic DNA, 20pmol of *C. elegans* GenePairs primers (Research Genetics) and 100μM dNTPs: 34 cycles of [94°C 30s, 58°C 30s, 72°C 90s] were followed by an extension of 1hr at 72°C to enhance A-tailing of products. Products were ligated into linearized T-tailed L4440 vector<sup>7</sup> and transformed into the HT115(DE3) bacterial strain (L. Timmons and A. Fire, pers. comm.) using standard methods. Colonies containing correct sized insert were identified by PCR using vector specific oligos, and the cloned inserts confirmed by PCR using the original Research Genetics primer pair. Primer sequences are available at <a href="http://cmgm.stanford.edu/~kimlab/primers.12-22-99.html">http://cmgm.stanford.edu/~kimlab/primers.12-22-99.html</a>.

RNAi screening. RNAi was performed essentially as described in Kamath et al8. where feeding data on 86 of the 2445 genes described here was previously reported. In brief, 4 wells of a 12-well plate containing NGM agar + 1mM IPTG + 25µg/ml carbenicillin were inoculated with bacterial cultures grown 8-18 hours for each targeted gene. 10-15 L3-L4 stage worms were placed in the first of the 4 wells for each gene and left for 72hrs at 15°C. Three worms, now young adults, were removed and individually placed on three remaining wells for each gene and allowed to lay embryos for 24hrs at room temperature; the three worms were then removed (t=0). The phenotypes of adults and progeny remaining in the first well were scored as well as of the progeny in wells 1-3. Our screen was not ideal for detection of phenotypes visible only in adults (e. g. egg-laying defective and progeny sterile); we will have missed some of these. Phenotypic analysis of lethality/sterility was carried out at t=24hr and post-embryonic phenotypes were analysed by two independent observers at t=36hr, t=48hr, t=60hr and t=72hr. Phenotypic classes were defined as follows. Embryonic lethal (Emb) reproducibly has 10-100% embryonic lethality; sterile (Ste) has a brood size of less than or equal to 10 (wild-type worms in these conditions typically give

over 50); progeny sterile (Stp) has a brood size of less than or equal to 10 in the progeny of fed worms. Post-embryonic phenotypes require at least 10% of the analysed worms to display a given phenotype; phenotypic classes are given in Table 1 legend. A full listing of phenotypes obtained is given in Supplementary Table 1; genes that we did not clone, and thus did not analyse, are given in Supplementary Table 3. Thus, any GenePair absent from both Supplementary Tables 1 and 3 was fed and did not give a detectable mutant phenotype.

## Bioinformatic analyses and categorisation of genes into functional classes.

Analyses were carried out on GenePairs predictions rather than currently predicted genes since while gene predictions change, phenotypes will always match the GenePair. ~95% of GenePairs genes have a one-to-one match with a currently predicted gene. Current gene predictions that are targeted for RNAi by the primer pairs were identified by comparing electronic PCR (ePCR) fragments (generated using the ePCR program (ftp.ncbi.nlm.nih.gov/pub/schuler/e-PCR)<sup>31</sup> on the whole chromosome DNA files from the WS9 release of ACeDB (ftp.sanger.ac.uk/pub/wormbase)) to gene predictions in ACeDB. To identify additional genes that might be targeted for RNAi by a particular clone we found those with an overlap of 200bp or more with greater than 80% nucleotide identity with the predicted PCR product (asterisks in column 2 of Table 3 denote GenePairs that have such matches); however it is not yet known what level of identity is required for RNAi.

To find *C. elegans* genes with conservation in other organisms, BlastP<sup>23</sup> was carried out for each individual *C. elegans* gene on Chromosome I against *S. cerevisiae*, *Drosophila melanogaster* and human sequences. The databases used were as follows: *C. elegans* (18337 entries), *S. cerevisae* (6191 entries) and *D. melanogaster* (13743 entries) downloaded on 1 June 2000 from www.ebi.ac.uk/proteome; and *H. sapiens* (35723 entries, confirmed peptides) downloaded on 1 June 2000 from

# Systematic functional genomic analysis of *C. elegans* Chromosome I by RNA interference

Andrew G. Fraser\*<sup>1</sup>, Ravi S. Kamath\*<sup>1</sup>, Peder Zipperlen\*<sup>1</sup>, Maruxa Martinez-Campos\*, Marc Sohrmann† and Julie Ahringer\*.

\*Wellcome/CRC Institute, University of Cambridge, Tennis Court Road, CB2 1QR Cambridge, UK. Phone +44-1223-334144; Fax +44-1223-334089. †The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

¹These authors contributed equally to this work.

Complete genomic sequence is known for two multicellular eukaryotes, the nematode Caenorhabditis elegans and the fruit fly Drosophila melanogaster, and will be soon for humans. However, biological function has been assigned to only a small proportion of the predicted genes in any animal. We used RNAmediated interference (RNAi) to target nearly 90% of predicted genes on C. elegans Chromosome I by feeding worms with bacteria that express double stranded RNA. We have assigned function to 13.9% of the genes analysed, increasing the number of sequenced genes with known phenotypes on Chromosome I from 70 to 378. While most genes with sterile or embryonic lethal RNAi phenotypes are involved in basal cell metabolism, many genes giving post-embryonic phenotypes have conserved but unknown function. In addition, conserved genes are significantly more likely to have an RNAi phenotype than genes with no conservation. We have constructed a reusable library of bacterial clones that permits unlimited future RNAi screens, which should help develop a more complete view of the relationships between the genome, gene function, and the environment.

The complete genomic sequence of an organism is an invaluable tool in understanding the molecular mechanisms underlying its development and function. The nematode worm *C. elegans* is one of two multicellular eukaryotes for which essentially complete genomic sequence is known<sup>1,2</sup>. 36% of predicted *C. elegans* genes have a significant human match<sup>1,3</sup> including many genes implicated in human diseases<sup>3,4</sup>, and functional analysis of the *C. elegans* genome has shed light on many conserved biological processes and molecular pathways. A comprehensive functional analysis of all genes in *C. elegans* would greatly expand our knowledge of conserved gene function. We therefore decided to investigate systematically loss-of-function phenotypes of predicted genes of *C. elegans*, starting with Chromosome I.

RNA-mediated interference (RNAi) is a technique whereby the activity of a gene is transiently inhibited following the introduction of double-stranded RNA (dsRNA) of sequence specific to the targeted gene<sup>5</sup>. The specificity and potency of RNAi make it ideal for investigating gene function beginning only with genomic sequence<sup>6</sup>. Ingestion of dsRNA-expressing bacteria results in RNAi of the targeted gene<sup>7</sup>, and we previously established that this technique is at least as effective as the injection of dsRNA for RNAi<sup>8</sup>: embryonic lethal phenotypes are detected with similar efficiency by feeding and injection, but feeding detects over 50% more postembryonic phenotypes than injection. It is thus possible to make a library of dsRNA-expressing bacteria which could be used for high-throughput genome-wide RNAi screens at very low cost. It is important to note that since RNAi does not efficiently inhibit all genes, an RNAi-based screen will miss some relevant genes. Despite this caveat, RNAi is a useful screening tool to complement classical forward genetics.

## Analysis of the function of genes on Chromosome I by RNAi

We constructed a library of bacteria expressing dsRNA corresponding to genes on Chromosome I. Chromosome I is the second smallest chromosome, has few

duplicated gene clusters and has no striking unusual features<sup>1</sup>. Each individual bacterial clone is able to synthesise dsRNA designed to target a single gene; since gene predictions are still changing, a few primer pairs no longer correspond to single genes (see Methods). In total, the resulting library contains 2445 independent clones, corresponding to 2416 predicted genes, a total of 87.3% of the 2769 currently predicted genes of Chromosome I.

We screened the library to identify genes whose inhibition gives a clearly detectable phenotype in wild-type worms as described in Methods. We were able to assign a phenotype to 13.9% of the analysed genes, raising the number of sequenced genes on chromosome I with known phenotypes from 70 to 378 (Table 1). Many genes have more than one associated phenotype, reflecting that genes frequently have multiple functions in the organism. Furthermore, since we examined worms that were only exposed to dsRNA as larvae or adults as well as their progeny, we could assign postembryonic phenotypes to genes that result in sterility or produce 100% embryonic lethal progeny. A summary of these results and a partial listing of the phenotypes obtained are given in Tables 1 and 3. Full results are in Supplementary Table 1 and arepublicly accessible in WormBase (www.wormbase.org).

Our screen was sufficiently effective to identify 90% of known embryonic lethal genes. In addition, we were able to assign phenotypes to 45% of genes with a known post-embryonic phenotype that should have been detectable in our screen (Table 2 and Supplementary Table 2). However, we failed to find phenotypes for a number of previously characterised genes. In some cases (e. g. *fog-3*), this was not due to an inherent difficulty in inhibiting the genes using RNAi (since we obtained the correct phenotype in a separate experiment), but simply because we overlooked them in the screen. However, only one of eight genes involved in neuronal function gave a detectable RNAi phenotype; this accords well with our finding that neurons appear to be more resistant to RNAi than other cell types<sup>8</sup>. Similarly, we did not detect

phenotypes for several genes involved in sperm development (fer-1, spe-9, and spe-11).

The largest phenotypic class, comprising over 60% of the genes, are those whose inhibition by RNAi gives rise to embryonic lethality, the Emb genes; these include a large number of components of the basal cellular machinery. More interestingly, we find a homologue of the SMN human disease gene<sup>9</sup>, a variety of genes encoding RNA-binding proteins (several such proteins play a role in early polarity; reviewed in <sup>10</sup>), a number of genes involved in chromosome condensation and separation, components of signal transduction pathways and many conserved genes that have no known biochemical function.

The largest class of post-embryonic phenotype is the Uncoordinated (Unc) class. Unc phenotypes arise from defects in the development or function of the neuromuscular system (reviewed in <sup>11</sup>). We find Unc genes encoding proteins involved in vesicle sorting and fusion as well as transcription factors (including a homologue of the zinc finger transcription factor MYT-1 which is only expressed in developing neurons in mammals<sup>12-14</sup>) and components of the cytoskeleton (e. g. a kakapo<sup>15-18</sup> and a talin<sup>19</sup> homologue).

A number of genes showed a high incidence of males (Him) phenotype. *C. elegans* is usually grown as a self-fertilising hermaphrodite with males arising at a low frequency in wild-type cultures due to non-disjunction of the X-chromosome (hermaphrodites have two X chromosomes, males only one). An increased number of males is indicative of either the incorrect segregation and maintenance of chromosomes in the germ line (reviewed in <sup>20</sup>) or defects in sexual specification. The Him genes that we identified include kinesins, a katanin homologue<sup>21,22</sup> and a nuclear hormone receptor.

### Conservation of genes with RNAi phenotypes across eukaryotes

We examined the level of cross-species conservation of the genes for which we detected an RNAi phenotype (Fig 1). To find *C. elegans* genes that are conserved in other species, we identified *C. elegans* genes that have hits with BlastP<sup>23</sup> e-values below 1.00E-06 in *Saccharomyces cerevisiae*, *Drosophila melanogaster* or humans; we define these as a "match". Hits with BlastP e-values below 1.00E-10 and in which the conservation extends over at least 80% of the *C. elegans* protein length, we defined as "homologues"; this category includes orthologues. This provides a conservative estimate of the number of genes with regions of conservation (matches) or homologues, respectively.

We found that genes with RNAi phenotypes were much more likely to have a match (p<0.001) compared to all genes (Fig 1). Most striking is the similarity that we see between *C. elegans* and *Drosophila*: while 42% of *C. elegans* genes have a match and 19% have a homologue in *Drosophila*, we find that over 72% of genes with an RNAi phenotype have a *Drosophila* match and 43% have a homologue (Fig 1). This analysis shows that genes with a required function in *C. elegans* have been highly conserved across eukaryotic evolution. We also find that highly conserved genes are more likely to have an RNAi phenotype than genes that show no conservation: 26% of *C. elegans* genes that have a homologue in one of the organisms examined give an RNAi phenotype compared to only 5% of genes with no conservation (p<0.001).

## Physical distribution on chromosome I of genes with RNAi phenotypes

Genes for which we identified an RNAi phenotype are evenly distributed across the chromosome with the exception of two regions (corresponding to segments 2 and 8-9 in Fig 2a) for which there appears to be a drop in number (p<0.1). These two regions correspond to the two regions of chromosome I that contain locally duplicated gene clusters<sup>1</sup>. We suggest that the reduction in the number of phenotypes observed by RNAi in these regions may be due to gene duplication and thus redundancy of function. It is worth noting that some of the predicted genes in the

duplicated regions may not be expressed: while genes with RNAi phenotypes are equally likely to have an EST in all regions of the genome (see below), there is a significant drop (p<0.05) in the proportion of total genes with ESTs in the second locally duplicated gene cluster region (Fig 2b; 39% of genes in the second cluster have an EST compared with 53% over the entire chromosome). We suggest that a portion of the predicted genes in such regions of duplication may in fact be pseudogenes.

Genes that give RNAi phenotypes are much more likely to have an EST than genes on chromosome I in general (82% versus 53% respectively, p<0.001; Fig 2b). The relatively high percentage of genes with RNAi phenotypes that have ESTs may reflect that these genes are expressed at higher levels. It may also be that many genes that currently lack ESTs are only expressed conditionally; we are unlikely to have found phenotypes for such genes.

In *C. elegans*, there is evidence of differences between the chromosome arms and the central regions (the clusters), suggesting that there might be differences in gene type or function across the chromosome<sup>24</sup>. In general, the distribution of genes in any given phenotypic class was similar to that for all genes with an RNAi phenotype (e. g. Emb genes; compare Fig 2c with 2a). However, genes with viable post-embryonic phenotypes (Pep genes) — those that gave a post-embryonic phenotype without any embryonic or post-embryonic lethality, sterility, or developmental delay — show a trend toward enrichment at the arms of chromosome I (p<0.1). It has been suggested that the chromosome arms may be more prone to mutation and recombination than the central core portion<sup>24</sup> and, if so, that novel gene functions are more likely to evolve in such regions. Our finding that genes which uniquely affect post-embryonic development cluster at the arms supports this model.

Relationships between the predicted biochemical function of a gene product and its RNAi phenotype

To explore the relationship between the biochemical function of a gene product and its mutant phenotype, we categorised the sterile (Ste), embryonic lethal (Emb), uncoordinated (Unc) and viable post-embryonic phenotype (Pep) genes into the functional classes shown in Fig 3a.

Unsurprisingly, genes involved in basal metabolic processes account for ~50% of Ste and Emb genes (Fig. 3a); this confirms that these basic biochemical processes are indeed essential for viability. In contrast, under 20% of Unc and Pep genes encode components of the basal metabolic machinery, whereas more than twice as many encode proteins with more specialized functions (Figs. 3a, b). There is thus a clear difference between the types of gene required for germline function or embryonic viability (which mainly require basal machinery) and those involved in later developmental processes which appear to require proteins either of more specialized functions or of as yet unknown function (Fig. 3b).

A second clear trend is that the number of genes of unknown function increases greatly in the Unc and Pep genes, making this the largest overall class for those phenotypes (Fig. 3). This shift underlies the fact that while we know a great deal about basic metabolic processes of eukaryotic cells (and thus can readily ascribe function to a large proportion of Ste and Emb genes), much is still to be learnt about the complex processes and the genes that regulate the development and function of a multicellular eukaryote. A significant number (~25%) of genes of unknown function have close homologues in *Drosophila* or humans; further study of these may shed light on conserved processes specific to animals.

## Comparison of genes essential for viability of S. cerevisiae and C. elegans

S. cerevisiae was the first eukaryote to be completely sequenced<sup>25</sup> and reverse genetics has been used extensively to investigate S. cerevisiae gene function. In a set of 3680 genes knocked out by targeted disruption, 890 affect viability<sup>26</sup>; we compared

these genes to those that gave different RNAi phenotypes in *C. elegans*. Yeast and worm genes important for viability have a similar distribution within the different functional classes, but are different from the Unc or Pep distributions (Fig 3c; also compare to 3a and 3b). This suggests that similar types of gene are required for viability of yeast and animal cells. A striking difference (p<0.001) is that only ~1% of the genes required for viability in yeast are transcription factors, whereas for *C. elegans* it is ~4% (a similar percentage of the genomes of yeast<sup>27</sup> and *C. elegans*<sup>2</sup> encode transcription factors, 3.3% and 2.5% respectively). This suggests that a large fraction of the *C. elegans* transcription factors required for viability may be involved in specific developmental processes.

## An estimate of the size of the functionally non-redundant genome

What do our data tell us about the size of the functionally non-redundant genome? We screened 12.7% of the *C. elegans* genome and found that 339 genes gave a clearly discernible phenotype. Taking into account the sensitivity of our screen and scaling up to the entire genome, we estimate that ~5400 genes will be individually required for wild-type *C. elegans* development under standard laboratory conditions (~2300 genes for embryonic viability and ~3100 post-embryonically; see Methods for calculation). This is comparable to previous estimates based on forward genetics<sup>28</sup>. We expect that phenotypes for other genes will be identified under novel conditions (e. g. environmental stress), in other genetic backgrounds, or using more refined and restricted screening conditions.

#### Discussion

We have taken a systematic approach to identify functions for the predicted genes of *C. elegans* Chromosome I. This is the first large-scale reverse genetic analysis of a multicellular organism and has increased by five-fold the number of sequenced genes with known phenotypes on this chromosome.

While we have identified RNAi phenotypes for many genes, some will have eluded our screen for one of at least two reasons. Firstly, RNAi may have been ineffective against the targeted gene. RNAi does not accurately phenocopy the null phenotype of all genes (e. g. genes involved in neuronal function), and may result in either partial or no loss of function. It should also be noted that if multiple genes have regions of identical or near-identical nucleotide sequence, RNAi could target them simultaneously, so that the observed phenotype may be the result of the inhibition of more than one gene. Secondly, we will not have detected either subtle or conditional phenotypes. However, we anticipate that future RNAi-based screens using specific assays should be able to detect phenotypes for many more genes, thus increasing our understanding of C. elegans and hence of metazoan biology in general. Since our library consists of bacterial clones that can be replicated, and the feeding protocol is relatively simple compared with injection, the library can be used repeatedly at low cost and high efficiency for such screens. In addition, we expect that a feeding library and database of associated phenotypes will prove valuable for the positional cloning of genes; currently there are over 300 genes on chromosome I identified by mutation but not yet cloned.

Although the time needed for an RNAi screen using our bacterial library is similar to that for a classical genetic screen, the two approaches have different advantages and will yield different results. Both approaches can be used to screen the entire genome for genes involved in a particular process, and both may identify complete or partial loss-of-function phenotypes. Classical forward genetics generates stable mutant lines that can be maintained indefinitely; furthermore, while some genes are resistant to RNAi, all genes are sensitive to mutagens (albeit to a greater or lesser degree) and could thus be cloned using a classical screen. Also, some mutants isolated by forward genetics are due to gain-of-function mutations, which cannot be generated by RNAi. However, the positional cloning of a gene is often slow and laborious. RNAi, while having the disadvantages mentioned above, has the key advantage of all

reverse genetics: the sequence of the gene is already known, and thus any mutant phenotype observed is automatically connected to a known sequence.

In the future, we aim to extend our library construction and functional analysis to the entire *C. elegans* genome and anticipate that the possibility of genome-wide RNAi screening, in conjunction with other functional genomics approaches such as expression analyses using microarrays<sup>29</sup> and two-hybrid experiments<sup>30</sup> will accelerate *C. elegans* research.

#### Methods

Generation and cloning of PCR products. PCR products were synthesised using BioTaq polymerase (Bioline) in a reaction containing 25ng of *C. elegans* genomic DNA, 20pmol of *C. elegans* GenePairs primers (Research Genetics) and 100μM dNTPs: 34 cycles of [94°C 30s, 58°C 30s, 72°C 90s] were followed by an extension of 1hr at 72°C to enhance A-tailing of products. Products were ligated into linearized T-tailed L4440 vector<sup>7</sup> and transformed into the HT115(DE3) bacterial strain (L. Timmons and A. Fire, pers. comm.) using standard methods. Colonies containing correct sized insert were identified by PCR using vector specific oligos, and the cloned inserts confirmed by PCR using the original Research Genetics primer pair. Primer sequences are available at <a href="http://cmgm.stanford.edu/~kimlab/primers.12-22-99.html">http://cmgm.stanford.edu/~kimlab/primers.12-22-99.html</a>.

**RNAi screening.** RNAi was performed essentially as described in Kamath *et al* $^8$ , where feeding data on 86 of the 2445 genes described here was previously reported. In brief, 4 wells of a 12-well plate containing NGM agar + 1mM IPTG + 25µg/ml carbenicillin were inoculated with bacterial cultures grown 8-18 hours for each targeted gene. 10-15 L3-L4 stage worms were placed in the first of the 4 wells for each gene and left for 72hrs at 15°C. Three worms, now young adults, were removed and individually placed on three remaining wells for each gene and allowed to lay embryos for 24hrs at room temperature; the three worms were then removed (t=0). The phenotypes of adults and progeny remaining in the first well were scored as well as of the progeny in wells 1-3. Our screen was not ideal for detection of phenotypes visible only in adults (e. g. egg-laying defective and progeny sterile); we will have missed some of these. Phenotypic analysis of lethality/sterility was carried out at t=24hr and post-embryonic phenotypes were analysed by two independent observers at t=36hr, t=48hr, t=60hr and t=72hr. Phenotypic classes were defined as follows. Embryonic lethal (Emb) reproducibly has 10-100% embryonic lethality; sterile (Ste) has a brood size of less than or equal to 10 (wild-type worms in these conditions typically give

over 50); progeny sterile (Stp) has a brood size of less than or equal to 10 in the progeny of fed worms. Post-embryonic phenotypes require at least 10% of the analysed worms to display a given phenotype; phenotypic classes are given in Table 1 legend. A full listing of phenotypes obtained is given in Supplementary Table 1; genes that we did not clone, and thus did not analyse, are given in Supplementary Table 3. Thus, any GenePair absent from both Supplementary Tables 1 and 3 was fed and did not give a detectable mutant phenotype.

## Bioinformatic analyses and categorisation of genes into functional classes.

Analyses were carried out on GenePairs predictions rather than currently predicted genes since while gene predictions change, phenotypes will always match the GenePair. ~95% of GenePairs genes have a one-to-one match with a currently predicted gene. Current gene predictions that are targeted for RNAi by the primer pairs were identified by comparing electronic PCR (ePCR) fragments (generated using the ePCR program (ftp.ncbi.nlm.nih.gov/pub/schuler/e-PCR)<sup>31</sup> on the whole chromosome DNA files from the WS9 release of ACeDB (ftp.sanger.ac.uk/pub/wormbase)) to gene predictions in ACeDB. To identify additional genes that might be targeted for RNAi by a particular clone we found those with an overlap of 200bp or more with greater than 80% nucleotide identity with the predicted PCR product (asterisks in column 2 of Table 3 denote GenePairs that have such matches); however it is not yet known what level of identity is required for RNAi.

To find *C. elegans* genes with conservation in other organisms, BlastP<sup>23</sup> was carried out for each individual *C. elegans* gene on Chromosome I against *S. cerevisiae*, *Drosophila melanogaster* and human sequences. The databases used were as follows: *C. elegans* (18337 entries), *S. cerevisae* (6191 entries) and *D. melanogaster* (13743 entries) downloaded on 1 June 2000 from www.ebi.ac.uk/proteome; and *H. sapiens* (35723 entries, confirmed peptides) downloaded on 1 June 2000 from

www.ensembl.org. NCBI-Blast2 was used (BLASTP 2.0.6) with the SEG filter, and the search space was set to 7947758.

We defined "sequenced genes with a known phenotype" as being those with a named entry in ACeDB that also have a known phenotype entered in ACeDB, WormBase (www.wormbase.org) or the Proteome database (www.proteome.com). EST data was supplied by the Sanger Centre on 21 June 2000.

Predicted gene products were placed into functional classes by manual inspection, primarily using data from Proteome, InterPro and Blast analysis<sup>23,32</sup>. The functional classes are: (1) DNA synthesis; (2) RNA synthesis and processing including general transcription machinery, splicing/processing, RNA binding and regulation of chromatin; (3) Protein synthesis and proteolysis including translation, degradation and folding; (4) Metabolism including energy production and intermediary metabolism; (5) Cell cycle and chromosome dynamics; (6) Cell biology and cellular structure including cell junction/adhesion, cytoskeleton, ion channels, protein trafficking and vesicle regulation and cell polarity; (7) Gene specific transcription; and (8) Signal transduction including kinases, phosphatases and components of signal transduction pathways. The Unknown functional class contains genes which either have motifs about which there is insufficient information to assign a function, or genes with no significant matches in any organism.

Estimates of non-redundant genome size were done as follows. We detected 90.5% of genes known to give an embryonic lethal phenotype and 32.6% of genes known to give a post-embryonic phenotype. After screening 87.3% of the genes on chromosome I, we identified 226 Emb genes and 113 genes that only gave a post-embryonic RNAi phenotype (including steriles); adjusting for our efficiencies of detection, we estimate that on chromosome I, 286 genes should be required for viability and ~397 for post-embryonic processes. We screened 12.7% of the genome,

and thus for the entire genome we expect 2250 Emb genes and 3130 genes to have a post-embryonic phenotype.

#### References

- 1. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012-2018 (1998).
- 2. Adams, M. D. et al. The genome sequence of *Drosophila melanogaster*. Science **287**, 2185-2195 (2000).
- 3. Rubin, G. M. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204-2215 (2000).
- 4. Ahringer, J. Turn to the worm! *Curr Opin Genet Dev* **7**, 410-415 (1997).
- 5. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811 (1998).
- 6. Tabara, H., Grishok, A. & Mello, C. C. RNAi in *C. elegans*: soaking in the genome sequence. *Science* **282**, 430-431 (1998).
- 7. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).
- 8. Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *C. elegans. submitted* (2000).
- 9. Lefebvre, S. *et al.* Identification and characterization of a spinal muscular atrophy- determining gene. *Cell* **80**, 155-165 (1995).

- 10. Bowerman, B. & Shelton, C. A. Cell polarity in the early

  Caenorhabditis elegans embryo. Curr Opin Genet Dev 9, 390-395 (1999).
- 11. Bargmann, C. I. Neurobiology of the *Caenorhabditis elegans* genome. *Science* **282**, 2028-2033 (1998).
- 12. Kim, J. G. & Hudson, L. D. Novel member of the zinc finger superfamily: A C2-HC finger that recognizes a glia-specific gene. *Mol Cell Biol* **12**, 5632-5639 (1992).
- 13. Armstrong, R. C., Kim, J. G. & Hudson, L. D. Expression of myelin transcription factor I (MyTI), a "zinc-finger" DNA- binding protein, in developing oligodendrocytes. *Glia* **14**, 303-321 (1995).
- 14. Kim, J. G. *et al.* Myelin transcription factor 1 (Myt1) of the oligodendrocyte lineage, along with a closely related CCHC zinc finger, is expressed in developing neurons in the mammalian central nervous system. *J Neurosci Res* **50**, 272-290 (1997).
- 15. Strumpf, D. & Volk, T. Kakapo, a novel cytoskeletal-associated protein is essential for the restricted localization of the neuregulin-like factor, vein, at the muscle-tendon junction site. *J Cell Biol* **143**, 1259-1270 (1998).
- 16. Gregory, S. L. & Brown, N. H. *kakapo*, a gene required for adhesion between and within cell layers in *Drosophila*, encodes a large cytoskeletal linker protein related to plectin and dystrophin. *J Cell Biol* **143**, 1271-1282 (1998).
- 17. Lee, S., Harris, K. L., Whitington, P. M. & Kolodziej, P. A. *short stop* is allelic to *kakapo*, and encodes rod-like cytoskeletal- associated proteins required for axon extension. *J Neurosci* **20**, 1096-1108 (2000).

- 18. Prokop, A., Uhler, J., Roote, J. & Bate, M. The *kakapo* mutation affects terminal arborization and central dendritic sprouting of *Drosophila* motorneurons. *J Cell Biol* **143**, 1283-1294 (1998).
- 19. Burridge, K. & Connell, L. Talin: a cytoskeletal component concentrated in adhesion plaques and other sites of actin-membrane interaction. *Cell Motil* **3**, 405-417 (1983).
- 20. Hodgkin, J. Sex, cell death, and the genome of *C. elegans. Cell* **98**, 277-280 (1999).
- 21. McNally, F. J. & Vale, R. D. Identification of katanin, an ATPase that severs and disassembles stable microtubules. *Cell* **75**, 419-429 (1993).
- 22. Srayko, M., Buster, D. W., Bazirgan, O. A., McNally, F. J. & Mains, P. E. MEI-1/MEI-2 katanin-like microtubule severing activity is required for *Caenorhabditis elegans* meiosis. *Genes Dev* **14**, 1072-1084 (2000).
- 23. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
- 24. Barnes, T. M., Kohara, Y., Coulson, A. & Hekimi, S. Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. *Genetics* **141**, 159-179 (1995).
- 25. Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 546, 563-547 (1996).
- 26. Mewes, H. W. *et al.* MIPS: a database for genomes and protein sequences. *Nucleic Acids Res* **28**, 37-40 (2000).

- 27. Svetlov, V. V. & Cooper, T. G. Review: compilation and characteristics of dedicated transcription factors in *Saccharomyces cerevisiae*. *Yeast* **11**, 1439-1484 (1995).
- 28. Hodgkin, J. & Herman, R. K. Changing styles in *C. elegans* genetics. *Trends Genet* **14**, 352-357 (1998).
- 29. Chen, J. J. *et al.* Profiling expression patterns and isolating differentially expressed genes by cDNA microarray system with colorimetry detection. *Genomics* **51**, 313-324 (1998).
- 30. Walhout, A. J. *et al.* Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science* **287**, 116-122 (2000).
- 31. Schuler, G. D. Sequence mapping by electronic PCR. *Genome Res* **7**, 541-550 (1997).
- 32. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

## Acknowledgements

We would like to thank Andy Fire and Lisa Timmons for kindly providing protocols, feeding vectors, and the HT115(DE3) bacterial strain and thank Richard Durbin, Daniel St.Johnston and Bjoern Schumacher for critical reading of the manuscript. A. G. F. was supported by a U. S. Army Breast Cancer Research Fellowship, R. S. K. by a Howard Hughes Medical Institute Predoctoral Fellowship, M. M.-C. by an EC-TMR Network Grant, P. Z. by a Wellcome Trust Studentship, M. S. by a Swiss National Science Foundation fellowship and J. A. by a Wellcome Trust Senior Research Fellowship (054523).

Correspondence and requests for materials should be addressed to J.A. (email: jaa@mole.bio.cam.ac.uk).

## Figure Legends

## Table 1 Summary of phenotypes arising from RNAi of genes on

Chromosome I. The number of predicted genes whose targeting via RNAi gave rise to each phenotype is shown. Percentages are given as percentage of total number of clones screened (2445). Phenotypic classes were defined as described in Methods. The phenotypes are Emb (embryonic lethal), Ste (sterile), Stp (sterile progeny), Gro (slow postembryonic growth), Lva (larval arrest), Lvl (larval lethality), Unc (uncoordinated), Pvl (protruding vulva), Bmd (body morphological defects), Dpy (dumpy), Clr (clear), Him (high incidence of males), Rup (ruptured), Mlt (molt defects), Prz (paralyzed), Sma (small), Egl (egg-laying defective), Sck (sick), Bli (blistering of cuticle), Muv (multivulva), Rol (roller), Adl (adult lethal), Lon (long), and Hya (hyperactive).

## Table 2 Detection of forward genetic loci on Chromosome I by RNAi.

RNAi phenotypes were compared to those of genes that have known loss-of-function phenotypes. "Genetic loci fed" denotes the number of genes in each category that were analysed by RNAi. "Possible to detect" denotes the number of genes that have a loss-of-function phenotype that would have been detectable in our screen. "RNAi phenotype detected" gives the number of genes for which a phenotype was identified. "Published phenotype detected" gives the number of genes for which the RNAi phenotype matched a published phenotype. Supplementary Table 2 gives full data. RNAi could reduce both maternal gene activity in the P0 and zygotic gene activity in the F1; this could explain some of the differences between RNAi phenotypes and published phenotypes.

Table 3 Partial list of RNAi phenotypes of genes on Chromosome I. RNAi phenotypes are shown for genes in the following functional classes: Chromosome dynamics and cell cycle; Cell structure; Specific transcription; and Signal transduction. For each gene, the following data are shown: the Research Genetics GenePairs name; whether a number of paralogues might be targeted (asterisk in column 2; methods gives criterion); the corresponding genetic locus name if it exists; a short description of gene function; the RNAi phenotype in which embryonic lethality (Emb), fecundity (Ste), post-embryonic phenotypes (P1-3) and developmental delay (Dev) are shown separately. Emb and Ste are classified into weak (white box, black "+") or strong (black box, white "+") phenotypes. For Emb, weak is 10-80% embryonic lethality, strong is 90% embryonic lethal or more; weak Ste denotes a brood size of 1 to 10, whereas strong Ste is totally sterile. Column H shows whether there is a match (white box, black "+") or a homologue (black box, white "+") in Drosophila melanogaster, Saccharomyces cerevisiae or humans. Phenotypic abbreviations are given in legend to Table 1. The GenePairs name does not always correspond with the current predicted gene name since gene predictions change.

Figure 1 Conservation of genes with an RNAi phenotype. Matches or homologues of *C. elegans* genes were identified as described in the text. Percentages of all genes (blue bars) or genes with RNAi phenotypes (red bars) with matches or homologues in *S. cerevisiae* (SC), *D. melanogaster* (DM), humans (HS), all three combined (ALL), or with no matches in any organism (NO M) are shown. The significance of the differences between the percentages of genes and the percentages of genes with RNAi phenotypes that have homologues is p<0.001 for all cases except for comparison to human homologues for which it is p<0.1.

Figure 2 Distribution on Chromosome I of genes with RNAi phenotypes and genes with ESTs. In each panel, chromosome I was analysed in 10 consecutive portions, each containing 10% of predicted genes. a) the percentage of all genes with an RNAi phenotype that are in each portion. b) the percentage of all predicted genes that have an EST (blue bars) or of genes that gave an RNAi phenotype that have an EST (red bars) in each portion. c) the percentage of Emb genes (black bars) or genes with viable post-embryonic phenotypes (pink bars) in each chromosomal portion. The boxes labelled "dup region" show the approximate location of regions containing local duplications.

Figure 3 Functional classes of Emb, Ste, Unc and Pep genes. Predicted products of genes that gave Ste, Emb, Unc or viable post-embryonic (Pep) RNAi phenotypes were placed into functional classes as described in Methods. Genes whose products could not be accurately classified into any of the 8 functional classes were placed into the unknown category (white). Numbers denote the percentage of genes in each functional class; pie charts illustrate these numbers graphically. b) Pie charts show distributions of predicted gene products grouped as follows: basal metabolic category (red) comprises the classes of DNA, RNA, protein and intermediate metabolism; specialized functions (blue) comprises cell cycle and chromosome dynamics, cell biology and cellular structure, gene specific transcription factors and signal transduction. Worms show the tissue affected in each phenotypic class shaded in grey. c) Distribution of genes giving rise to non-viable RNAi phenotypes in *C. elegans* (worm) or to non-viable phenotypes following disruption in *S. cerevisiae* (yeast).

Supplementary Table 1 Phenotypes arising from RNAi of genes on

Chromosome I. Genes that have a detectable RNAi phenotype are grouped

by the functional classes shown in Fig 3. For each gene, the following data are shown: the Research Genetics GenePairs name; whether the sequence might target a number of paralogues (asterisk in column 2; methods gives criterion for this); the corresponding genetic locus name if it exists: a short description of gene function; the RNAi phenotype in which embryonic lethality (Emb), fecundity (Ste), post-embryonic phenotypes (P1-3) and developmental delay (Dev) are shown separately; existence of matches (lower case) or homologues (filled box, white upper-case text) in C. elegans (CE), Drosophila melanogaster (DM), Saccharomyces cerevisiae (SC) or humans (HS); and whether or not the gene has an EST (E). Abbreviations used are Ste (sterile), 1-5 (fed worm had 1-5 progeny), 6-10 (fed worm had 6-10 progeny), Stp (progeny sterile), LvI (larval lethality), Unc (uncoordinated). PvI (protruding vulva), Bmd (body morphological defects), Dpy (dumpy), Clr (clear), Him (high incidence of males), Rup (ruptured), Mlt (molt defects), Prz (paralyzed), Sma (small), Egl (egg-laying defective), Sck (sick), Bli (blistering of cuticle), Muv (multivulva), Rol (roller), Adl (adult lethal), Lon (long), Hya (hyperactive), Gro (slow post-embryonic growth) and Lva (larval arrest). "Mult" indicates that the gene has multiple equalpenetrance post-embryonic phenotypes. If the dsRNA overlaps multiple adjacent genes of different function, these appear in the "Multiple genes" category. The GenePairs name does not always correspond with the current predicted gene name since gene predictions change.

. . . .

Supplementary Table 2 RNAi phenotypes for previously identified loci on Chromosome I. Columns 1 and 2 give genes on chromosome I with previously identified embryonic lethal or post-embryonic phenotypes and the GenePairs primer pair that amplifies a fragment overlapping that gene, respectively. "Mutant Phenotype" gives the published phenotype. RNAi phenotype headings: "Emb" (percentage embryonic lethality); "Ste" (sterility);

"P1", "P2" and "P3" (post-embryonic phenotypes); and "Dev" (slow or arrested growth). "Hit" indicates whether an RNAi phenotype was obtained in the initial screen (tick), whether no mutant phenotype was obtained ("o") or whether a mutant phenotype was obtained in separate feeding experiment ("\*"). Phenotype abbreviations are given in the Supplementary Figure 1 legend with the following additions: Slu (sluggish), Vul (vulvaless), Mec (mechanosensory abnormality), Daf (dauer larva formation abnormal), Ttx (thermotaxis abnormal), Che (chemotaxis defective). Also, the following abbrevations are used: phen (phenotype), migr (migration), red (reduced), wk (weak), abnl (abnormal), and dk (dark). Genes with null phenotypes that we would have failed to detect in our screen are shaded in light grey; genes that we failed to clone, and therefore failed to analyse, are shaded in dark grey.

Supplementary Table 3 GenePairs on Chromosome I for which no clone was obtained.

Table 1 Summary of phenotypes arising from RNAi of genes on Chromosome I

PHENOTYPE		NUMBER	PERCENT
All phenotypes	TOTAL	339	13.9
Embryonic lethal	Emb	226	9.2
Sterile	Ste	82	3.4
	Stp	14	0.6
Developmental delay	Gro/Lva	145	5.9
Larval lethal	Lvl	38	1.6
Post-embryonic	Unc	70	2.9
	Pvl	29	1.2
	Bmd	27	1.1
	Dpy	19	0.8
	Clr	14	0.6
	Him	13	<b>0</b> .5
	Rup	9	0.4
	Mlt	8	0.3
	Prz	8	0.3
	Sma	6	0.2
	Egl	5	0.2
	Sck	5	0.2
	Bli	4	0.2
	Muv	2	0.1
	Rol	2	0.1
	Adl	1	< 0.1
	Lon	1	< 0.1
	Hya	1	< 0.1

Table 2 Detection of forward genetic loci on Chromosome I by RNAi.

I MOIO Z D COCCOOL OI IOI			•	
Phenotype	Genetic loci fed	Possible to detect	RNAi phenotype detected	Published phenotype detected
All phenotypes	62	50	31	25
Embryonic lethal	21	21	19	16
Sterile Sterile progeny	3 4	3 4	2 1	2 1
Developmental delay Larval lethal	0 4	0 4	- 1	- 1
Post-embryonic	43	31	14	9

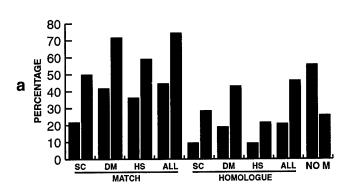
Table 3 Partial list of RNAi phenotypes of genes on Chromosome I

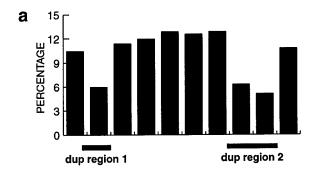
. . .

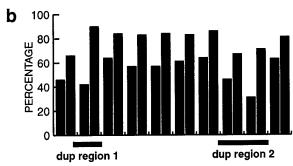
GenePairs		Locus	Description	Emb	Ste	P1	P2	Р3	Dev		GenePairs		Locus	Description	Emb	Ste	P1	P2	P3	Dev	H
, c											Cell struc	ture	•								
B0511.8			CDC1-like	+	-	-	-	-	Lva	+	Y34D9A_152.a			vacuolar sorting	+		Unc	Prz	LvI	Gro	+
C37A2.4		cve-1		+	一	Clr	-	-	-	+	Y48G8A_3945.e	П		adaptin subunit	-	-	Unc	-	-	•	+
C41G7.2	٠	<u> </u>	kinesin	+	-	Him		-	Ι-	+	Y71A12B.a	П		gravin-like		-	Dpy	Clr	•	•	+
C53H9.2			chrom stability	+	-	-	-	-	Gro		Y71F9A_279.b	П		NXT1 homol	+	-	PvI	Unc	-	-	+
F57B10.12		mei-2	katanin homol		-			-	-	+	Y71F9A 282.b	П		coatomer subunit		+	Unc	-	-	•	+
M01E11.6	•	11101-2	kinesin	+	-	Him	-	-	Gro		Y71F9A_290.a	П		NTF2 homol	-	-	Pvl	Rup	Clr	-	+
R06C7.8			Bub1-like	+	-	PvI	Rup	LvI		+	Y71G12A_195.e	П		talin	+	+	Unc	Prz	-	-	+
T01G9.5		mei-1	ATPase	+	-	-		-	-	+	Y87G2A.s	П		HuVPS28 homol	-	-	Unc	-	-	-	+
W09G3.3	٠	11101	RCC1 domains	-	-	Him	-	-	-	+	Y87G2A.x	П		protein trafficking	•	-	Unc	-	-	Gro	+
Y110A7A.d	_		cdc27 homologue	+	-	-		-	Gro	+	Y87G2A.y	П		protein trafficking		-	Clr	-	-	-	+
Y39G10A_246.e			MCM4-like	+	-	-	-	-	-	+	ZK1014.1	П		vesicle fusion	-	+	Lvi	•		-	+
Y39G10A_246.i	-		INCENP-like		-	-	_	-	-	+	ZK1151.2	$\vdash$		spectrin repeats	-	-	Unc	Bmd	Stp		+
Y47G6A_247.i			pombe Rad2 homol		Ť		-	Ι-	-	+	EKT TO 1.2			<u> </u>							_
Y52B11A.9			Kin17 homol	+	-	Bmd		-	-	+	B0025.3			txnl corepressor	+		-		-	Gro	+
Cell struct	117		KIIII / HOHIOI	+		Dillu	nup				C01G8.7	Н		evelid-like	+		Bmd	LvI	-	Gro	
	uic		Ezrin-like		Ι	Unc	LvI		Gro		Ç01G8.8	$\vdash$		eyelid-like	+	-	Bmd	-	-	-	+
C01G8.5			villin	H	-	Him		÷	-	+	C01H6.5	Н	nhr-23	nuc horm recep.	-	_	Unc	LvI	Dpy	-	+
C10H11.1				+	+	Unc	_	÷	-	Ť	C12C8.3	Н	lin-41	NHL domains	-	+	-	-		-	+
C17E4.9	_		PDZ domain	-		-	-	÷	÷	+	C32F10.7	╁╌	nhr-2	nuc horm recep.	-	-	Him	-	-		+
C32E8.10		unc-11	vesicle reg	+	÷	-		÷	-		C48E7.3	Н	11111-2	bZIP	-		-	-		Gro	
C45G3.1	_		actin-binding		$\overline{}$	Unc				+		╁╌┤		SRF homol	-	-	Unc	Prz	-	-	+
C47B2.3	_	tba-2	tubulin		-		•	-	•	+	D1081.2	Н		TCF-9-like	- +	÷	-	F12		Gro	
C53D5.a	_		nuclear import		+	-	Ŀ	٠	Gro	+	F43G9.12	Н			•	÷	Unc	-	-1	•	-
C53D5.i			nuclear import		•	-	٠	-	Gro	+	F52F12.6	$\vdash$		MYT1 homol		-	-			Lva	+
DY3.2		lam-1	nuclear lamin	+	+	LvI		Ŀ	-	+	F55F8.4	Н		txnl repression			_		-		
E03H4.8	•		beta coatomer-like	-	+	Unc	-	Cir	Gro	+	F57B10.1	Н		bZIP	+	+	Sma	Dpy	$\overline{}$	Lva	
F07A5.7		unc-15		-	٠.	_	Prz			+	K02B12.1	Н	ceh-6	homeobox	-	•	Unc	Mit		٠	+
F20G4.3		nmy-2			+	-	-	-		+	M05B5.5	Н	hlh-2	bHLH		+	Unc		-	•	+
F21C3.5			MT nucleation	+	-		Bmd		Gro	+	W02D3.9	Н		groucho family	+	+	Unc		•	-	+
F26B1.3			karyopherin		-	Lvi	-	-	<u> </u>	+	W03D8.4	Н	pop-1	HMG box	+	+	-		-	-	+
F26E4.8	٠	tba-2	tubulin	+	-	-	<u> </u>	-	<u> </u>	+	Y40B1A.4	Н		Zn finger	-	•		Bmd	<u>.</u>		+
F26H9.6			ras superfamily	+	+	LvI	-	-	<u> </u>	+	Y54E5B.3	Ш		Mediator complex	+		Rup		-	•	+
F28H1.2			calponin domain	+	<u> </u>	-	-	-	Lva	+	Y65B4A_179.b	Ш		txnl activator	-	-		Dpy	•	-	+
F30A10.6			transporter	-	-	<u> </u>		-	Gro	+	ZC247.3			LIM homeodomain	-		-	-	-	Gro	
F36H2.1			cation transporter	+	-	-	<u> </u>	-	Gro	_	ZK858.4		mel-26	kruppel-like	+	•	<u> </u>	٠		-	+
F43G9.10			microfib assoc	+	-	-	-	-	Gro	+	Signalli	ng									,
F46F11.5			vacuolar ATPase		+	LvI		-	ŀ	+	C09D1.1			multiple domains	<u> </u>	-	Unc	-		-	+
F53B8.1	*		plectrins		-	Unc	Prz	Lvl	-	+	C10H11.9			ROCK	+	<u> </u>	Dpy	Rol	LvI		+
F53F10.5			nucleoporin-like	+	-	-	-	-	-	+	C26C6.2	Ш	goa-1		+	+	Unc	PvI	Egl	-	
F54C1.7			troponin c	-	+	•	-	-	١-	+	C32E8.5			FHA domain	+	+	-	•	-	Gro	
F55A12.7			UNC-101 homol	-	-	Unc	٠	-	Gro	+	F26E4.1	L		PP2A reg subunit	+	•	Lvl	-	-	-	+
F55F8.5			MT associated	-	-	-	٠	-	Lva	+	F55A12.3	Ш		PI-4P 5' kinase	-	+	<u> </u>	-	-	-	+
F56F4.5			transporter	-	-	Him	-		-	+	F55C7.4		unc-73		+	•	-	-	-	•	+
H15N14.1			human NSF-like	+	+	Ŀ	-	-	-	+	F55C7.7	$\Box$	unc-73		-	-	Egl	-	-	-	+
M01A10.3			ribophorin			LvI	Unc	-	-	+	K04G2.8	Ĺ		APC homol	-	-	Unc	Bmd			+
R05D11.3			NTF2 homol	+	+	-	-	-	Gro	+	K05C4.6			B-catenin	+		Unc	Dpy	Bmd		+
T03F1.9	_		UNC-89-like	+	-	-	-	-	-		K12C11.2	П		SUMO-1 like	+	-	PvI		-	-	+
T19B4.2			NUP153-like		+	-	-	-	-	+	T01G9.6		kin-10	CKIIbeta subunit	+	-	Pvi	-	-	Gro	+
T21E12.4	_	dhc-1	dynein heavy chain		-	-	-	-	-	+	T21E3.1	•		PTPase	+	+	٠	-		-	+
T25G3.2			chitin synthases	+	-	-	-	-	Gro		T23D8.1		mom-5	frizzled-like	-		Unc	Bmd		-	+
T26E3.3		par-6	PDZ domain		-	-	-	-	-	+	T23H2.5	•		rab family	-	-	Sma	Cr	-	-	+
W02B9.1		hmr-1	cadherin	+	-	Bmd	Unc	Dpy	-	+	Y106G6E.6	П		Casein Kinase 1	+	-	Dpy	•	-	Gro	+
W04C9.1		,,,,,,,,,,	ABC transporter	+	-	-	-	-	Gro	+	Y18D10A.5		gsk-3	GSK-3	+		-	-	•	-	+
Y105E8C.n	-		gamma-adaptin	7.	_	Unc	_	-	Gro	+	ZC581.1	Т	4	NIMA-like kinase	-	-	Unc	LvI	Rol	-	+
Y18D10A.17	-		sup of clathrin defic	_	-	-	-	-	-	+	ZK265.6	П		G prot coup recep	-	-	-	-	-	Gro	+
Y18D10A.17	-	nfn-1	profilin	+	Ť	<del>  -</del>	-	-	† <del>.</del>	Ť				L							
1100104.20		Pin-1	promit						٠	_											

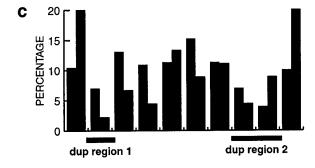
<sup>'</sup> Figure 1

Fraser *et al.* 

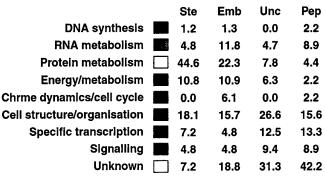


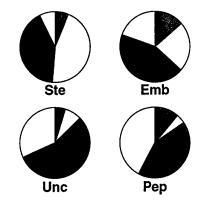


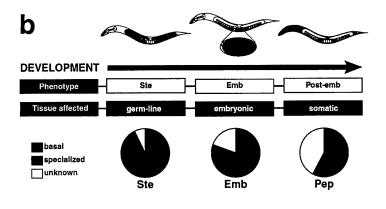


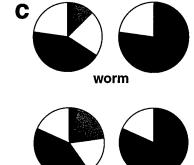


	١
v	
_	









yeast

ONA primage (ONA pot appra authorn)   ONA primage (ONA pot appra authorn)   ONA primage (ONA pot alpha authorn)   ONA pot alpha on primage (ONA pot alpha on primage)   ONA pot alpha on primage (ONA pot alpha on primage)   ONA pot alpha on primage (ONA pot alpha on primage)   ONA pot alpha on primage (ONA pot alpha on primage)   ONA pot alpha on primage (ONA pot alpha on primage)   ONA pot alpha on primage on properties   ONA pot alpha on probate on pr	M01E5.5		0	A pol epidon subunit	100%	,	•	-	•		j	į	ī
Other prices (The Lot after a school)   100%   50   10   10   10   10   10   10   10			٥				Mon	윤			Ę	Ž.	+
High gradies   1975	W02D9.1		0	-	8		Σ	4		-		9	+
Michael Control Cont	ZC308.1		٥	i	90%	8	•	4	ş	-	ŧ	-	+
High protein   High protein   1979,   1970	Chapte 2	1	- 1	SMA cercina an terms	-	Γ.	A PLANE	5	-	8		-	
Harking protein   Harking	C17E4 5	3	1	N-A hinding profesion	806		2	3,	5	8		+	•
Physical description but   1507   100	2608.1	L	Ξ	AG protein	-	Γ.	Σ	ļ.	8		_	2	+
Note the content of	38085	L	2	stone descetylase comolex subunit			A V	E E	Ŀ		ş	٤	+
Second potential formation and the second potenti	C36B1.3	L	Œ	AA pol II subunit	100%			_	ΘB		Ę	4	+
Micharden grieding price in the 1-STAG GLD 1   00%   10%	C36B1.5		4	licing factor	100%	Se	Σ	-	g	8	 C	2	
State of the production   1000   10	C37A2.1		Œ	A-binding protein like T-STAR, GLD-1		٠	ž	1	-	8	Ę	-	+
March   Marc	1.751.1		ō	AN orthologue	ŝ		8	1	3	Ī	ı		+
Marche   March   Mar	Z48E7.2	1	I	A por III subunit			•	1	5			٠,	٠
Control Cont	2000	1	8	Varive splicing lactor	8	·	+	1	5	8	į	2	•
March   Marc	818010	1	2 1	VO-HKB, TONG RI SCHOOL	ŝ	·	. 2	1	1			2	
Coltant National Health Commission   1909;   170;	F2643.5	1	-	close mDNA can hinding	200	Ţ	2		5			20	
Style*Cetting and Activation   100%	E2643.3	8	-	Ī	2000			ŀ	ě	8	-	ŀ	ŀ
Cities RNApoli (1007)   Citi	F30F8 8		-		,00 100%				Ŀ	I.	B	2	+
Committee   Committee and total Committee   Committee   Committee   Committee   Committee   Committee   Committee   Committee and total Committee   Committee   Committee   Committee   Committee and total Committee   Committee and total Committee   Committee and total Committee   Committee and total Committee   Comm	F32H2.1		S	xeldu	%08-0%		-	Ŀ	ફ		튱	8	+
The control of the	F36F2.3	L	0	tike RNApol II CTD phosphatase	100%		•	_	Ŀ		Ę	2	+
Protection   100%   1	F37E3.1		Ε	RNA CAP binding	100%		일	Ė	g		.io	8	+
mass   Net formatin ingulation   100%   150	F49D11.1		ā	tative splicing factor, WD-repeat	%		Š		•	3	20	E C	+
The control protection of explosion   100%	F53G12.5	¥9E	¥ وم	1 domain	100g		-	-	•		ŧ	-	+
Characteristic protein services   100%   1	Ī		?	volved in chromatin regulation	. 1000		8	1	1			ł	+
Chicked Chickey   Chicke	1	ļ	5 0	A serial direction actor	8		. 8	+	. 6	1	į	ŀ	
The control of the	13/DIU.0		5	NO porthologies	Š		₹.	1	1	20	,	ŀ	•
The control interesting   100%   PM   LM   C   C   C   C   C   C   C   C   C	K02B12.3	L	15	inacriotional repressor		ŝ	+	1	g		통	1	
The 1 chromatic menceduling   100%   PM   Unc   C   C   C   C   C   C   C   C   C	K02F2.3	L	55	licino factor	100%	,	Σ	Ľ	ļ.	35	Dir.	-	+
The Act of Promodeling   100%   PM   LM   C   C   C   C   C   C   C   C   C	K07A1.11	ē	-	romatin remodelling	100%		<u>ح</u>	8	•	픙	MO	30	+
Network HA	K07A1.12	đ	2	romatin remodelling	200%		3	2	Ŀ	3C	NO	90	+
Particular of the control of the c	T23D8.5		Z	stone H4	10%	٠			ဋ	8	통	26	
PACE   RINALAMENT   Francisco   100%   100	T23D8.6				%			-	•	ö	Ę	20	+
The processing of the proces	T23G11.3	Š			800	ĝ			٠	8	Ę	8	+
The Properties about 1	A01811.3		*		% 04 02		•		Š	ö	ö	-	+
Welving Companies B, spicing rich   100%   1-5 Ray	WO4A8.7		۲	AF II complex subunit TAF2A	20%		,		•		ě	Ľ	+
High Coll   Recomplication   100%   10%   10%   100%   1	WD8E3.1	_	*	RNP polypeptide B; eplicing role	100%	5	Ş		٠		ŧ	8	+
RNA pot II abburt   100%   510   0   0   0   0   0   0   0   0   0	110A7A.m			licing component	100%	٠		-	•	8	ć	8	+
HINP G   Except   100%   6-10   100%   6-1	7G8A 247.h		~	VA pol III subunit	•	٠	ຮັ	4	-			1	+
Exercise   Problem   15   100%   510   100%   10%   100%   100%   100%   100%   100%   100%   100%   100%   100%	E10B 159.c		-	NA pol II subunit	8	-	1	-	4				•
Control protein List   100%   6-10   10%   6-10   100%   6-10   100%   6-10   100%   6-10   100%   6-10   100%   6-10   100%   6-10   100%   6-10   100%   6-10   100%	1F9B 286.b		•	RNP G	క్త	.]	Š	5 *					+
Control Cont	in matcheolism	-	1		/9000	5	-	-	-		Ī		[
Continue	BOE 1 10	1	2	preferious installer factor	ŝ	6	. 3	1	1	I	1	ı	
Chaptering protein; 124   10076, Sha   10076	mane 1		1	concernal pertain 124	8	3		Η.			ě		
Chaptering protein: L1.4   100%, State	00000	ļ	ľ	concernal prototol 1.24	1	ŀ	ŀ	ľ	d		è	1	
The control protein 1, 13   100%   Sie   1   1   1   1   1   1   1   1   1	CO4F12.4	ļ	1	posomal protein L14	300	8	ŀ	H.	Ş		ã	2	+
The continue of the continue	C09D4.5	L	Ī	osomal protein L19	%	2		ŀ.	3				+
Chrosenia protein existent   100% Ste LM   600 65 00	C32E8.2	L	Ē	bosomal protein L13	4001	ŝ		Ė	·				+
Translational protein   Control	C36B1.4	L	٩	otessome subunit	100%	g	3	Ė	ဗီ	S.			+
Introductional Holdson ledor   107%   144   6 do 0   07   55   15   15   15   15   15   15   1	C37A2.7	L	=	xxecmal protein			•	Ë	စ်			-	+
Required for 655 beganesis   100% See   100	C41D11.2		-	snstational initiation factor	100%		3	-	Š	_ 1	Ę	-	+
Proplement attachment   10076, See   100	C43E11.4		•	ongation factor Tu		$ \cdot $			ē		20	÷ S	+
Professore subunit   107% Sile   107%   10   10   10   10   10   10   10   1	C43E11.9		٤	quired for 60S biogenesis				-	ဒီ		NO	SC	+
Integrated for 165 processing   100%   15   10   10   10   10   10   10   10	C7884	4	٦	oteasome subunit	200%	ĝ	ŝ			8	MO	S	+
Incomment protein   10072, 1-5 St.   100 00 00 00 00 00 00 00 00 00 00 00 00	C4782.5		•	£		$\cdot$		-	ទ័		i d	SC	+
The control potation   100%   15   1   1   1   1   1   1   1   1	C48B6.2		-	quired for 18S processing			g	-	ğ		Š	9	1
Mark Agriculation   1077, 318   10   10   10   10   10   10   10	D1007.12		=	posomal protein	Š	2	•	:	3		Ž.	9	+ 0 0
HVV A part   HVV	97/0010			posomai protein	2	8	2	7	1	4		١	+
Accordance production   100%   50 to 10	122003	1	Ī	INA SYNDRIASE	3	2/8	•	1	ž	1	5 6	200	
## 17	1.23rg.10	1	1	Dosomai prosein	, 000	8		1	+		5	,,	+ -
The confidence of the confid	E SHE'S	-	+	Otessome audunin	38	8	5 2	2	+	3 8	Š	2 (	
	F25H5.4	É	7		ŝ	8	Ξ	1	,		5	2 8	+
1975   1975	F28H1.3		_		20-90%	?	•		5	-	10	S.	+
Ordering processes, carried and a continue of contin	F32A7.6	×	-		%	·	1	1	5	8	€.	8	+
Introduction integron reportmentage   10,775   Sec.   Co.   Co.	F32B5.8	4	1		Š			:	+	8	Ē		+
District Control Con	2011.10	1		anstational muation technicase	3	·	+	+		Š		36	0 0
Controlled   Con	Coop to	1	1		80.00	8	1	+	5	į	Š	3 (	
Committee   Comm	Carbon P	1	Ť		200	í				Ş	8	3	
Continued Operation   1975	COOL 14 F	-			8 8	8	3	-		1	1	3	
Comparison   Com	E4849.4	ļ	֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓		%08	,	2	2	ě		ΣĊ	Ş	
Command protein LS	C48A9 5	1	1		8	ŀ			ė	i i	à	8	
	CEOPE 9	1	,	min based degradation	8 9	. 8		+	5	38	i	3	,
Contraction power   Contraction   Contract	12783.0	1			8 8	8	·	+	. 6		i	36	0 0
	153612.10	4		2	3	8	·	,	5		Š	36	•
Milk processing   Milk processing   Milk ply LM   Milk processing   Milk ply LM   Mi	F56H1.4	1	٦		80	2	·	+	1		á	9	+
Mid-processing   Mid-py-LiM   1.05	F59A3.3	1	-	U-B	85	2	1	+	<u>د</u>		Ì	-	+
14   400-568   170-568   170-568   170-57   17	F5906.4	-	Т	NA processing					=+			8	
Trimington induction index   100%   50 mil Strain   100%   100 mil Strain   100 mil Strain   100%   100 mil Strain   100 mil Strain   100%   100 mil Strain   100%   100 mil Strain   100 mil Strain   100%   100 mil Strain   100 mil Stra	K04F10.4	Ż	- 1	ubikase protease						ő	ě	8	+
A A Titless pulsacrone suburif (1075, 5 or 0 or	K04G2.1	-	7	ensistional initiation factor	100%	2	2	2	- 1		٤	8	+
	K04G2.3	4	Ť	AA ATPase, proteasome subunit	20-40%	·	1	+	5	_	1	8	+
Dougland State of the Can projection of the Can projection of the Can c	K05C4.1	4	Ť	roteasome subuni	250	8		1	1		5 8	2 5	+ 1
protessome suburit	HOSES.4	1	Ť	Decum-conjugating	2000	٠		5	-	_	5	3	+
	R12E2.3	4	٥	the second section and second		έ		H	-	L	1	7	-

++	+	П	+ +	+ +	+ +	+	+ 1	+	+ +	+	1	Т	+	П	+ +	+	+	+	+	+	T.	+ +	+	+ 4	+	+	+	+	+	+	+ +	Π	+	++	+	1	$\cdot \Box$	+		+	+ +	+	+	+[	+	+	* *	П	+		+			Ш		
77	Ħ	П	2 2	+	2	M	+	П	t	H	Ħ	†	Ħ	Τ.	2	2		Ė	T	2			ž	2	2 2	2	2	2	T	2	2	П	1	2 2	2	T	18	П	1	2	ė,	뾷	П	1	П	2	T	П	T	П		П				ı
	$\dagger \dagger$	$^{\dagger}$	Н	$^{+}$		Г	+	Ħ		П		†	П	1	200	8			8	Н	1	Ī	8	8		T	8	8	T	T	T	Ħ	Ť	8	8	T	T	П	T	8	ò	0	П	T	T	1	T	П	T			П		П		ĺ
	5 8	Ę	Ę	+	3 2	a	+	Ħ	Ę	Ę		$\dagger$	Ħ		s c	Ē	Н	8	B	ŧ	ı	Ė	H	£ £	ı	ă.	П	통	T	Ę		H	Ē	ŧ	Æ	-	ć	T	Ť	Ę	5			T	П	ŧ	T		T	П		П	7	Т	П	١
		$\mathbb{H}$	1	$^{+}$	127	Ī	1	3	Ť	8		t		ľ	8	•	H	-	ľ	8	Ī		8	8	-		П	ď	8	Ī		8	Ť	8	,			П	T	П	Ī	Г	8	g	3	1	1	П	†	П		П	T	T	П	ı
S S		. g		۶.	δ 6	6			. 8	$\perp$	.1.	+	G G	٠,	2 5	Lva	1		. g	-	1	.†.	g	₹.	g	9	g	õ	. B	g	. 5	Π,	5 G		g	8 8			٠,			5		·	3 .	e i	9 8		. [				g.	Gro	,	ı
			. F			δ		3	Ĕ.	1			+	3		-	1	7		Pup			·	7		·	Ī	-		1		1		<u> 3</u>				•				1	٠	·				ı			3	1		,	·	ı
1 2		æ .	. ₹		1.				ğ .		2 .		T,	ž	٠,		ŀ	1		Doy.	1					•	1	7	· 8	1	, .			£ 8	1	•			٠.	•		•	•	Σ		٠	. Pr				Prz		. 4		·	1
. 5	ξĒ	₹ .	E S	울.	3	Sto Unc	. 8	£	ğ .	Ē	3 .	E S		<b>8</b>	ğ .		ŀ	7	. 3	Ş	Ę	<u> </u>	٠	. 3	À	ຮັ	3	•	Σ	1	5 6	5	Che of	200	8		3	. 3	5	ŏ.	g,	1	S D	Š	,	٠	8 5		3	} ·	Š		Ě	Bmd	,	
H	1-1	<b>8</b> .		Sle .	1.	١,	,	Ī.				Ţ	١.			Se Se		,		١.		. 10											Τ.		1.		,					,	,							See				Γ.	ŀ	
28.8	1 1	- 1	- 1 - 1	- 1	╁┼	38	28	+	$^{+}$	H		+	%	-	+	H	٥	او	ه ه	H	+	-	Н	8		%	٥	+	۰	8	8	H	. %	26	28	2 <sup>2</sup> 2		.e.	2 2	H	1		Н	+	×	+	+	8	+	26	Н	8	28 26	٠	×	
50-80%	}	20-90%	100% 50-80%	20-80		20-80%	8	808	1	. 8	38	1	20-90%	٠		100%	9	8	88	Ŀ	1	100%	+	20-40%	%08	2040	100%	. 8	<u>}</u>	50-80%	20-80%	. 8	20-40%	. 04	50-80%	8 8	100	90	3 8		100%	3	'	1	100%	1		20-80%	1	100°		100	50-80%	S	100	
and the Appendix A DAN domeins	Ispendite Of a liberia, 11 Parties		coiled coils DnaJ domain		colled coils		coiled coils	coiled coile	PF01942: domain of unknown function						WD repeat		colled colls	coiled coils		ankyrin repeat		elioo belloo	Dros homologue Peter Pan for larval grow		mago nashi functional orthologue		well-conserved domain at C-ten	GTPase	PDZ domains	coiled coils		BTB/POZ domain, large elegans family		DEAD-box binding TrkA-fused cene homol		coiled colls	CHORD-containing protein		COHEC COHE		collect colle	Anna cons					colled coils									
																			dad-1							kup-2																										Ц				
	Ш			- 1	H	Ш	4	+	4	Н	+	4	+	Н	+			+	$^{+}$	H	Н	t	H	H	H	Н	Н	+	Ė	+	+	H	+	+	H	+	+	H	$^{+}$	H	$^{+}$	H	H	1	t	+	+	Н	+		1		+	H	H	
	H	+		+	П	П		11	-	ш			i i		- [	1																							۰ اــ	L	_l_	1 1	اا	,	40	٦.	ام	11	- 1	11					اہا	ŀ
																								L	,					_ ,	_	۵,		E 40	, 0	١.	,	6	9	9	9	3	50	ıΩ	3 8	59	8 2	ő	3	ایران	_	<b>1</b> 2	۳,		4	
482.8	167.3	587.5	1A21 C12.2	004.1	308.6	5H2.4	8E4.4	704.2	2,110,9	13.2	2H2.3	342.6	6A2.7	G12.4	SF8.3	6A3.2	6A3.3	8A3.4	810.10	PA11.1	2B12.8	845.4	9.646	184.1	983.5	169.4	3011.4	946.2	165.4	3083	148.4	2A11.8	855	10D5.3	SEBC.d	6G6D.1	0A7A.g	3H5A.3	10A 246	6A 245 b	64 247.8	1A 54.d	31B 55.0	18 55	0A 156	10A 159.	4A 182 c	34B 10.c	74B 12.a	308.3	328.1	9H11.2	7A12.3	308.9	31A_54.t	
C3482.8	C4167.3	C55B7.5 D2030.3	E01A2.1 F18C12.2	F20G4.1	F23C8.6	F25H2.4	F26E4.4	F2704.2	F30A10.9	F31C3.2	F32H2.3	F33H2.6	F36A2.7	F53G12.4	FSSF8.3	F56A3.2	F56A3.3	F56A3.4	F57810.10	K02A11.1	K02B12.8	K06A5.4	K09H9.6	K1184.1	R0883.5	T01G9.4	T08G11.4	T19A6.2	12165.4	T23D8.3	WO1A8.4	W02A11.8	W09C5.4	W10D5.3 Y83D3A 5	Y105EBC.d	Y108G6D.1	Y110A7A.g	Y23H5A.3	Y39G10A 246	Y47G8A 245 b	Y47G6A 247.8	Y48G1A_54.d	Y48G1B_55.0	Y48G1B 55.	YS4E10A_156	Y54E10A 159.	Y6584A 182 0	Y65B4B_10.c	Y65B4B 12.8	20308.3	ZC328.1	F39H11.2	K07A12.3	T23D8.9	Y48G1A_54.	

100% Sie LM · · · Circ Dir Si · · · · Sociol Dir Si · · · · · · · · · · · · · · · · · ·	LM Gno dm 500 Sign Unc Prz LM Gm sc	Unc   OR   OR   S   OR   OR   OR   OR   OR	100% Sie Livi Unic · · · · Orm   Ns   100% Sie · · · · · · · · Orm   Ns   100% Sie · · · · · · · · · · · · · · · · · · ·	Se on one of the contract of t	69 8 8 9 9	Unc Prz Lvi Gro Die sie	· Unc · · · · · · · · · · · · · · · · · · ·	₹ § ₹	Prz	Sie CM · · · · · · · · · · · · · · · · · ·	Unc Bind Stp · 6# ex ec	g g .	888	88 E5 E5		Ste Sme Dpy · Lva dm · Unc Mft · · · os dm	100% Ste 100% 1-5	SO-80% - 1 Uno Dry - 0 os dm ec	Unc	88	2 · 6	8 8 8	Bmd Lvi	<b>3 5</b>	Sie	50-80% Dpy Gro CE DM SC HS 100%	. Unc LM Rol . CE DW SC	- Gro	Quo	Bmd PM	
	Merofibrillar AP homologue vacuolar, ATPasa 2 genes, both plectrins nucleopolin homology	troponin c clathrin associated; 64%ID UNC-101 clathrin associated oliopopidide transporter vesicular fusion; human NSF-like	https://doi.org/10/10/10/10/10/10/10/10/10/10/10/10/10/		tier Un AP-1		deptin complex subunit ike Hu gravin domain for localisation of PK	NI F2-related NX I 1 orthologue coatomer complex alpha subunit NTF2-related export protein	alin 16%1D to Hu VPS28	sec/ comain, non-dathin protein traffick sec/ domain, non-dathin protein traffick vesicle fusion	spectrin repeats; kakapo-like	corepressor of nuclear hormone receptors oyelid-like eyelid-like	Tuckear normone recepturi NHL domains Tuckear hormone receptor	AZIP transcription factor 55%ID to SRF	Hu ICF-9-like Zn finger, Hu MYT1-like ranacriptional repression	oZIP transcription factor nomeobox transcription factor NH Hamendrich factor	groucho family transcription factor HMG box protein	An imper transcription lactor Mediator complex autumit transcriptional activator Lan homeodomain transcription factor knuckel-like Zn finser transcription factor	multiple signalling domains	G-a subunit	PP2A regulatory subunit	guanine exchange factor	APC homologue	SUMO-1 like regulatory subunit of CKlibeta	protein tyrosine phosphatase	Casein Kinase 1	NIMA-like kinase G-protein coupled receptor			WD repeat	
			Ę			1-mg							in-41			Seh-6	pop-1	In-11	89-Jun	-Bog		unc-73a	apr-1	kin-10	mom-5	6,38					
F26H9.6 F26H1.2 F30A10.6 F36H2.1	369.10 6F11.5 5388.1 3F10.5	F54C1.7 F55A12.7 F55F8.5 F56F4.5 H15N14.1	5011.3 5011.3 351.9 1984.2	563.2 8653.3 8653.3	W04C9.1	D10A.20	8A_3945.e 1A12B.a	Y71F9A 279.b Y71F9A 282.b Y71F9A 290.a	12A_195.e	787G2A.y ZK1014.1	ZK1151.2	B0025.3 C01G8.7 C01G8.8	C12C8.3	48E7.3	F52F12.6 F55F8.4	F57B10.1 K02B12.1	0203.9 70308.4	Y4081A.4 Y54E58.3 Y65B4A_179.b ZC247.3 ZY858.4	Signelling C0901.1	C28C8.2	26E4.1	55C7 4	04G2.8	12C11.2 01G9.6	T21E3.1 T23D8.1	7106G6E.6	C581.1 7K265.6	Inknown B0261.1	B0511.9 C01A2.5	306AS.1	25A16

<b>EMBRYONIC</b>	LETHALS									
					RNAi	phenotype				Similar to
Locus	GenePairs Name	Mutant Phenotype	Emb	Ste	P1	P2	P3	Dev	Hit	published
apr-1	K04G2.8	Emb	-	•	Unc	Bmd	Lvl			0
dhc-1	T21E12.4	Emb	100%	-			•		/	
ego-1	F26A3.3	Ste, Emb	50-80%		<u> </u>	. <u>.</u>	•	Gro	/	/
hlh-2	M05B5.5	Emb	100%	6-10	Unc	PvI	-		/	/
hmp-2	K05C4.6	Emb	100%	-	Unc	Dpy	Bmd	- 1	/	/_/
hmr-1	W02B9.1	Emb	10%		Bmd	Unc	Dpy		/	/
let-502	C10H11.9	Emb	50-80%	-	Dpy	Rol	Lvl			/
mei-1	T01G9.5	Emb	100%			<u> </u>		- 1	/	/
mel-26	ZK858.4	Emb	100%	-		•		-		1
mex-3	F53G12.5	Emb	100%				-			/
mom-5	T23D8.1	Emb		•	Unc	Bmd	-	<u> </u>	_/	0
nhr-2	C32F10.6	Emb	-		Him	-				0
nmy-2	F20G4.3	Emb	50-80%	Ste		<u> </u>	•		/	/
par-6	T26E3.3	Emb	100%				•	-		/
pop-1	W03D8.4	Emb	100%	1-5	· .	-				/
rba-1	K07A1.11	Emb	100%		PvI	Unc	-		/	/
rba-2	K07A1.12	Emb	100%	•	PvI	Lvl		-	/	/
unc-37	W02D3.9	Unc, Emb, Ste	100%	Ste	Unc	-	-	-	/	/
unc-73	F55C7.2,3,4,7	Emb, Unc, Egl, Sma, Dpy	20-40%		Egl	-	-		/	/
mom-4	F52F12.3	Emb	-	•			-		0	0
sup-17	DY3.7	Emb, Dpy, Unc			<u> </u>	· .	-		0	0

3	GenePairs Name F32A7.6 K04F10.4 K02F2.2 T23G11.3 C26C6.2 ZC247.3	Mutant Phenotype constipated Bli Dpy, Lvl	emb 10% - 20-40%	Ste	P1 -	phenotype P2	P3	Gro
Locus aex-5 bli-4 dpy-14 gld-1 goa-1 lin-11 lin-11 mes-3 nhr-23	F32A7.6 K04F10.4 K02F2.2 T23G11.3 C26C6.2 ZC247.3	constipated Bli Dpy, Lvl	10%	-				
bli-4 dpy-14 gld-1 goa-1 lin-11 lrp-1 mes-3	K04F10.4 K02F2.2 T23G11.3 C26C6.2 ZC247.3	Bli Dpy, Lvl	-					Gro
dpy-14 gld-1 goa-1 lin-11 lrp-1 mes-3	K02F2.2 T23G11.3 C26C6.2 ZC247.3	Dpy, Lvl			Mit	Dov	Lvl	-
gld-1 goa-1 lin-11 lrp-1 mes-3	T23G11.3 C26C6.2 ZC247.3			-	Unc	Lon		Lva
goa-1 lin-11 lrp-1 mes-3	C26C6.2 ZC247.3		50-80%	Ste	-	-		
lin-11 Irp-1 mes-3	ZC247.3	Hya, Egl-c	10%	1-5	Unc	PVI	Egl	-
Irp-1 mes-3		Vul, wk Unc	10%	- 1-5	-	- 1 11		Gro
mes-3	E000111	Sma, Dpy, Mit		<del></del> -	Unc	Prz		Gro
	F29D11.1	Stra, Dpy, Will			Stp			
nnr-23 I	F54C1.3		<del>                                   </del>		Unc	LvI	Dpy	
	C01H6.5	Mit, Let, wk Emb	10%		- Oilc	- LVI	DDy	-
unc-11	C32E8.10	Unc, wk thin	10%		Unc	Prz	Egl	
unc-15	F07A5.7	Unc, Prz, Egl			Unc		- Egi	
unc-37	W02D3.9	Unc, Emb, Ste	100%	Ste -			<del></del>	
unc-73	F55C7.2,3,4,7	Emb, Unc, Egl, Sma, Dpy	20-40%		Egl	:_	-	
unc-89	C09D1.1	wk Unc, Clr, thin	-	-	Unc			
egl-30	M01D7.1,7	Slu, Egl	<u> </u>		-	-	<b>:</b>	
fer-1	F43G9.6	Ste, lays oocytes		-	-	-	<u> </u>	$\vdash$
fog-3	C03C11.2	Ste, no sperm	<u> </u>	-				
gsa-1	R06A10.2	Lvl		-	<u> </u>	:		
lin-10	C09H6.2	Vul, Muv	<u>                                    </u>	-	<u> </u>			
spe-11	F48C1.7	paternal-effect Emb	<u>                                     </u>	•		-	-	
spe-9	C17D12.6	Ste	<u> </u>	•	•	-	•	
sup-17	DY3.7	Emb, Dpy, Unc			<u> </u>	-	<u> </u>	
sur-2	F39B2.4	Let, Vul, Egl		-	<u> </u>	. •		<u> </u>
unc-101	K11D2.3	Unc, wk Egl	<u>                                     </u>		•	<b>-</b> _		
unc-13	ZK524.2	Unc, Prz		•	<u> </u>	-	<u> </u>	
unc-14	K10D3.2	Unc, Prz	<u>                                    </u>		<u> </u>			<del>   </del>
unc-29	T08G11.5	Slu L1	<u> </u>	•	<u> </u>		-	<u> </u>
unc-38	F21F3.5	Slu, wk Dpy	<u>                                    </u>	-	-		•	<del></del>
unc-40	T19B4.6,7	wk Unc,wk Dpy		_ •	<u> </u>	•	•	<u> </u>
unc-55	F55D12.4,6	slow Unc	l L	-	-	•	-	<u> </u>
unc-87	F08B6.4	Unc, Prz	-	Markey are with make	(III) and the Appellation (III)	SOLUTION OF STREET	Di Amiliana desa	-
		Unc. Prz						
air-2	not cloned not cloned	Emb. Ste wk Unc. Lon, wk Egl. Muv						
lin-17	not cloned	Egl						
mek-2	not cloned	Ste, Vul						
sing-2	not cloned	PvI						
spe-12	not cloned	Ste						
	not cloned	Ste						
spe-4 unc-54	not cloned not cloned	Unc						

GenePair	GenePair	GenePair	GenePair	GenePair	GenePair	GenePair	GenePair
B0019.3	E03H4.4	F56A6.4	T23H2.6	Y34D9A_150.a	Y54E10A_157.a	Y71G12A_205.f	ZC434.8
B0025.2	F02E9.2a	F56G4.5	T27C10.4	Y34D9A_150.b	Y54E10B_151.a	Y71G12A_205.g	ZK1053.1
B0041.7	F02E9.6	F57C9.2	T28F2.3	Y34D9A_151.c	Y54E10B_151.b	Y74C10A_140.a	ZK1053.2
B0205.10	F08B6.3	H16D19.3	VH15N14R.1	Y34D9A_151.f	Y54E10B_152.c	Y74C10A_140.b	ZK1053.4
B0205.6	F11C3.2	K02B12.4	W02D3.1	Y34D9A_152.b	Y54E10B_154.a	Y74C10A_148.a	ZK1151.3
B0205.7	F11C3.3	K02B12.8	W03D8.10	Y34D9A_152.c	Y54E10B_154.b	Y74C10A_150.a	ZK265.2
B0205.9	F12B6.3	K05C4.10	W03D8.7	Y39G10A_233.a	Y54E10B_159.d	Y74C10A_150.b	ZK39.4
B0207.1	F14B6.1	K07A1.10	W04C9.6	Y39G10A_241.a	Y54E5A.3	Y74C10A_151.a	ZK484.4
B0207.11	F14B6.2	K07A1.2	W05B5.1	Y39G10A_243.b	Y63D3A.4	Y74C10A_151.b	ZK524.1
B0207.4	F14B6.5	K07A1.5	W05B5.3	Y39G10A_244.a	Y65B4A_165.a	Y74C10A_152.a	ZK524.3
B0207.5	F15D3.7	K08C9.4	W05F2.1	Y39G10A_244.b	Y65B4A_185.b	Y74C10A_152.b	ZK524.4
B0261.2	F15D3.8	K10C3.5	W05H12.1	Y39G10A_246.g	Y65B4B_13.a	Y74C10A_153.a	ZK849.4
B0261.5	F21A9.1	K10E9.2	W09C5.4	Y40B1A.3	Y65B4B_15.d	Y74C10A_154.a	ZK858.3
B0261.7	F21A9.2	K11B4.2	W09C5.5	Y44E3A.2	Y65B4B 9.a	Y74C10A 154.b	ZK909.2
B0379 5	F22D6.8	K12C11.3	W09C5.6	Y47G6A 232.8	Y6B3B.6	Y74C10A 154.6	ZK909.5
BO414 6	F22G12.2	M01D7.2	W09G3.5	Y47G6A 232.b	Y71F9A 261.a	Y74C10A 156.a	ZK973.1
004116.0	0 01303	M01D7.4	14(1000 2	V4706A 240 A	V71E0A 260 2	V74C10A 156 h	7K973 m
01110.2	72013.2	MOLD1.4	W1000.2	747064 246.0	V7450 004 0	V74010A 157.0	77000 H
C04F12.1	F20H9.3	MUIG12.12	W1003.2	14/G0A 243.6	1/1F9A 201.8	1/40104 13/.8	ZN993.D
C04F12.2	F26H9.5	M01G12.3	Y105E8A.d	Y47G6A_246.d	Y71F9A_291.b	Y/4C10A_15/.b	
C04F12.5	F27C1.8	M01G12.5	Y105E8A.e	Y47G6A_247.d	Y71F9A_294.a	Y74C10A_158.a	
C04F12.6	F28B3.6	M04F3.2	Y105E8B.c	Y47G6A_247.j	Y71F9A_294.b	Y74C10A_158.b	
C10H11.4	F29D10.5	R05D7.4	Y105E8B.d	Y47G6A_247.k	Y71F9A_294.g	Y74C10A_158.c	
C15C6.1	F31C3.1	R06A10.4	Y105E8C.a	Y47H10A.a	Y71F9A_295.b	Y74C10A_158.d	
C17D12.1a	F32A7.4	R06C1.2	Y105E8C.b	Y47H10A.b	Y71F9B 275.a	Y74C10A 158.e	
C17H1.3	F32A7.7	R11A5.2	Y105E8C.c	Y47H9C.2	Y71F9B_276.b	Y74C10A 158.f	
C17H1 8	F32B4.2	R11A5.5	Y105E8C.d	Y47H9C.5	Y71F9B 287.c	Y74C10A 158.0	
C18E3.2	F32B4.5	R12E2.1	Y105E8C.h	Y48G10A.c	Y71F9B 297.b	Y87G2A.c	
C18E3.8	F32B4.8	R12E2.8	Y105E8C.m	Y48G10B.a	Y71F9B 299.b	Y87G2A.i	
C25A1.8	F32H2.2	T01A4.2	Y105E8C.u	Y48G1A 53.c	Y71G12A 187.c	Y87G2A.k	
C27C7 4	E32H2 9	T01H8.3	Y105EBD a	Y48G1A 54 a	V71G12A 189 a	Y87G2A m	
C27C7.8	E33D11 4	T02F1 1	V105F8D h	Y48G1A 54 B	Y71G12A 190 a	Y87G2A.0	
C31H5.6	F33D11 7	T02F1 5	V105ERD e	V48G1C 55 b	V71G12A 190 h	Y87G2A n	
23407	20000	TOCE	V105EBD #	V48G8A 2614 h	V71612A 101 2	V87G2A v	
3402.3	0.300.	10000	7400000	C40000	V34C12A 102	20000	
C3051.9	F30F2.4	10200.3	1100000.1	140G0A 2014.C	771012A 193.8	10/32A.X	
C3/A5.2	F3/D6.2	102G6.4	Y 106G6D.Z	148G8A 3051.C	1/1G12A_195.8	19568A 82.8	
C41G7.4	F39H2.2	T03F1.7	Y106G6E.5	Y48G8A_3304.a	Y71G12A_195.d	Y95B8A_82.b	
C48B6.5	F40E3.6	T04D1.3	Y110A7A.k	Y48G8A_3671.b	Y71G12A_196.a	Y95B8A_84.a	
C49A1.10	F43G9.8	T04D1.4	Y18D10A.15	Y48G8A_3671.c	Y71G12A_197.a	Y95B8A_84.e	
C49A1.4	F46A8.1	T05E7.5	Y18D10A.3	Y48G8A 3945.f	Y71G12A 199.a	Y95B8A, 85.b	
CERDE	EAGAR 10	T07010 5	V18H14 67 a	V48G8A 4850 a	V71G12A 199 h	V95B8A 85 d	
CESDE P	EAGAR A	TODE 11.3	V18H1A 67 h	V48G8A 5005 a	V71G12A 199 c	V9588A 85 f	
11.600.0	1,000.4	5.113601	110111 07.0	140004 2333.8	1,1915A 199.C	Leg Vocase	
C54C8.6	F4962.4	109E11.4	Y18H1A_6/.e	Y48G8A_5995.D	171G12A 202.b	19568A 85.9	
C54G6.1	F49D11.5	1.059.1	Y18H1A b/.n	Y52B11A.11	471G12A 203.D	ZU24/.1	
C54G6.3	F49D11.7	T10E9.5	Y18H1A_67.I	Y52B11A.3	Y71G12A_203.d	ZC247.4	
C55C2.1	F52A8.2	T15D6.1	Y18H1A_68.b	Y53H1A.a	Y71G12A_203.e	ZC334.1	
C55C2.3	F52F12.7	T15D6.4	Y18H1A_68.c	Y53H1A.c	Y71G12A_203.f	ZC434.1	
D1037.2	F53B6.1	T19A6.1	Y18H1A 68.e	Y53H1B.c	Y71G12A_205.a	ZC434.3	
D2005.6	F53F10.3	T22C1.1	Y23H5A.1	Y53H1C.c	Y71G12A 205.c	ZC434.4	
DV3.6	E54A5 3a	T22C1.3	Y26D4A.10	Y54E10A 156.b	Y71G12A 205.d	ZC434.5	